



Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

**Délivré par** *l'Institut National Polytechnique de Toulouse*  
**Discipline ou spécialité :** *Informatique*

---

**Présentée et soutenue par** *Pascaline PARISOT*  
**Le** *23 janvier 2009*

**Titre :**

*Suivi d'objets dans des séquences d'images de scènes déformables*  
----  
*De l'importance des points d'intérêt et du maillage 2D*

---

**JURY**

<i>Michel DHOME</i>	<i>(DR – président du jury)</i>	<i>LASMEA Clermont-Ferrand</i>
<i>Stéphane PATEUX</i>	<i>(IR)</i>	<i>France Télécom R&amp;D Rennes</i>
<i>Peter STURM</i>	<i>(DR)</i>	<i>INRIA Grenoble</i>
<i>Marie-Odile BERGER</i>	<i>(CR)</i>	<i>INRIA Lorraine</i>
<i>Alain AYACHE</i>	<i>(PR)</i>	<i>IRIT</i>
<i>Bernard THIESSE</i>	<i>(MdC)</i>	<i>IRIT</i>
<i>Géraldine MORIN</i>	<i>(MdC – invitée)</i>	<i>IRIT</i>
<i>Vincent CHARVILLAT</i>	<i>(MdC – invité)</i>	<i>IRIT</i>

---

**Ecole doctorale :** *Mathématiques-Informatique-Télécommunications de Toulouse*

**Unité de recherche :** *Institut de Recherche en Informatique de Toulouse*

**Directeur(s) de Thèse :** *Alain AYACHE*

**Rapporteurs :** *Michel DHOME, Stéphane PATEUX, Peter STURM*



<i>Titre :</i>
----------------

Suivi d'objets dans des séquences d'images de scènes déformables : de l'importance des points d'intérêt et du maillage 2D
---

<i>Résumé :</i>
-----------------

<p>Nous abordons le suivi d'objets dans des séquences d'images de scènes déformables selon deux axes de recherche. Il s'agit de déterminer les transformations d'un objet, d'une image à l'autre, lorsque celui-ci s'est éventuellement déformé ou déplacé et lorsque le point de vue de la caméra a éventuellement été modifié (déplacement, zoom...). Pour cela, nous nous sommes inspirés de l'algorithme de Jurie et Dhome qui permet de suivre un objet plan indéformable. D'une part, nous en améliorons les performances. D'autre part, nous le généralisons au cas d'objets déformables.</p>
--

<p>Le premier axe de recherche consiste à améliorer les performances de l'algorithme de Jurie et Dhome en termes de précision et robustesse. Le suivi s'appuie sur un ensemble de points d'intérêt, dont dépendent fortement les performances. Ces points d'intérêt sont issus d'une sélection des points obtenus par des détecteurs reconnus, à savoir SIFT, KLT, SUSAN, HARRIS et MORAVEC. Nous avons étudié et mis en œuvre, sur différentes classes d'images, des heuristiques de sélection fondées sur des approches statistique et algébrique. Nous montrons :</p>
--

- |  |
|--|
| <ul style="list-style-type: none"><li>– qu'il n'existe pas de détecteur universel,</li><li>– que l'approche statistique est à privilégier dans tous les cas.</li></ul> |
|--|

<p>Le second axe de recherche est une proposition d'un nouvel algorithme de suivi s'appuyant sur le maillage 2D des images de la séquence. Cet algorithme généralise celui de Jurie et Dhome aux scènes déformables. Il repose sur :</p>
--

- |   |
|---|
| <ul style="list-style-type: none"><li>– des transformations élémentaires (nodales) du maillage, directes et inverses, que nous avons caractérisées tant d'un point de vue géométrique qu'analytique,</li><li>– l'utilisation des coordonnées barycentriques généralisées pour approcher la composition de deux transformations d'un maillage.</li></ul> |
|---|

<p>Cet algorithme donne des résultats similaires à celui d'appariement hexagonal de Nakaya et Harashima tout en étant plus rapide.</p>
--

<i>Mots Clés :</i>
--------------------

suivi, apprentissage, points d'intérêt, maillage 2D, objets déformables, séquence d'images
--

<i>Title :</i>
----------------

Objects tracking in video of non rigid scenes : importance of interest points and 2D mesh
---

<i>Abstract :</i>
-------------------

<p>We deal with object tracking in videos of non-rigid scenes with two main purposes. We aim at determining the transformations of an object, from one frame to the next, when it may be distorted or moved and when the camera focus may change (movement, zoom...). To do this, we were inspired by the Jurie and Dhome algorithm, which enables the tracking of plane rigid objects. On the one hand, we improve its performance. On the other hand, we generalize it to non-rigid objects.</p>
--

<p>The first goal consists in improving the performance of the Jurie and Dhome algorithm, in terms of accuracy and robustness. The tracking is based on a set of interest points, which has a great effect on the algorithm's performance. These interest points come from a selection among the points extracted with some common detectors: SIFT, KLT, SUSAN, HARRIS, and MORAVEC. With various pictures classes, we have studied and implemented some selection heuristics based on statistical or algebraic approaches. We show that:</p>
---

- |  |
|--|
| <ul style="list-style-type: none"><li>• there is no universal detector,</li><li>• the statistical approach is the best in all cases.</li></ul> |
|--|

<p>The second goal is a proposal of a new tracking algorithm based on a 2D mesh of the video frames. This algorithm generalizes the Jurie and Dhome one for non-rigid scenes. It is based on :</p>
--

- |  |
|--|
| <ul style="list-style-type: none"><li>• elementary (nodal), direct or inverse, mesh transformations that we geometrically and analytically characterize,</li><li>• generalized barycentric coordinates to approximate the composition of two mesh transformations.</li></ul> |
|--|

<p>This algorithm gives similar results to the hexagonal matching algorithm of Nakaya and Harashima while being faster.</p>
---

<i>Keywords :</i>
-------------------

tracking, learning, interest points, 2D mesh, non rigid objects, video
--

# **MÉMOIRE**

de

**Pascaline PARISOT**

intitulé

## **Suivi d'objets dans des séquences d'images de scènes déformables**

—

## **De l'importance des points d'intérêt et du maillage 2D**

Composition du jury :

---

<i>Rapporteurs :</i>	Michel DHOME Stéphane PATEUX Peter STURM
<i>Examineur :</i>	Marie-Odile BERGER
<i>Responsables :</i>	Alain AYACHE Vincent CHARVILLAT Géraldine MORIN Bernard THIESSE

---



À Maman,





# Remerciements

Il est de coutume d'avoir une page de remerciements dans tout mémoire que l'on écrit... mais qui vais-je donc pouvoir remercier ? Après réflexion, j'ai trouvé quelques personnes qu'il me tenait à cœur de remercier plus particulièrement...

Je tiens tout d'abord à remercier Monsieur Alain Ayache, professeur à l'ENSEEIH, pour m'avoir donné la possibilité d'effectuer une thèse au sein de son équipe.

Mes remerciements vont également aux membres de mon jury Madame Marie-Odile Berger et Messieurs Michel Dhome, Stéphane Pateux et Peter Sturm pour m'avoir fait l'honneur d'évaluer mes travaux de recherche. Merci à ces trois derniers pour leurs remarques, en tant que rapporteurs, sur ce mémoire.

Mes sincères remerciements à mes responsables Géraldine Morin, Vincent Charvillat et Bernard Thiesse pour m'avoir guidée dans cette recherche, pour les discussions que nous avons eues ensemble. Plus particulièrement merci à Bernard pour la rédaction de ce mémoire, pour sa tenacité, pour son perfectionnisme... Merci à vous trois pour tout ce que j'ai acquis en vous côtoyant.

Merci à mes relecteurs pour le travail de longue haleine qu'ils ont effectué : Céline, Philippe, Noële, Sandrine et Sébastien.

Je tiens également à remercier mes co-bureaux successifs pour m'avoir supportée... dur, dur... Christophe, Pascal, Bernard, le Viok (heu... Stéphane) et Benoît.

Merci aussi à tous les doctorants ou stagiaires avec qui j'ai parcouru un bout de chemin et passé de très bons moments : Matthijs, Ariel, Vincent, Christophe, Bertrand, Aurélie, Cezar, Jean-Charles, Stéphane (et le pingouin volant), Sébastien, Sandrine (et les coin-coin lumineux), Nassima, Victoria, Benoît, Jérôme, Dorin... Merci pour leur sympathie, leur disponibilité, leurs plaisanteries.

Merci à tous les membres de l'équipe VORTEX pour m'avoir très bien accueillie pendant ma thèse : Jano et sa constante bonne humeur, Romulus et ses pots improvisés, Pierre, Jean-Pierre, Philippe...

Un grand merci à toutes les personnes que j'ai côtoyées durant mes enseignements et qui ont été confrontées à mes innombrables questions. Merci pour votre patience et vos réponses. Un remerciement tout particulier à Xavier C. pour sa rigueur, sa clarté d'explications et les échanges fort enrichissants que j'ai eus avec lui. Merci. Merci à Marc, Xavier C., Vincent, Géraldine, Bernard, Pascal, François, Ronan, Marcel, Christiane, Pierre, Daniel, Aurélie, Sandrine, Nassima... pour m'avoir fait confiance et m'avoir permis d'enseigner aussi bien l'informatique que les mathématiques appliquées.

Sans oublier tous les collègues du laboratoire, voisins de bureau pendant un temps, animateurs de discussion, tous sympathiques : Xavier T., Emmanuel, André-Luc, Julien, Jérôme, Béatrice, Gérard, Michel...

Une pensée aussi pour le personnel du secrétariat et du support technique que j'ai eu le plaisir d'embêter pendant cette thèse pour leur sympathie et leur patience : Annie,

Anne-Marie, Violette, Josiane, Sylvie E., Sabyne, Sylvie A., Marie-Blandine, Jean-Claude, Stéphane.

Je voudrais aussi remercier des personnes qui restent souvent dans l'ombre et qui font qu'aujourd'hui j'en suis là. Je pense à tous les enseignants que j'ai eus, qui aimaient leur métier et qui m'ont transmis leur savoir et leur passion pour leur matière. Merci donc à Mesdames Borel, Dumas, Viallet... et à Messieurs Castelli, Cournède, Dumas, Queyrel... Tous ceux-là et bien d'autres ont joué un rôle important dans mon parcours scolaire qui prend fin avec ce mémoire.

Enfin, je tiens à remercier les proches, famille et amis, qui m'ont accompagnée pendant ces années.

Tout d'abord, merci à mon amie Céline pour son soutien tout au long de mon cursus à l'ENSEEIH et de la thèse, nos innombrables discussions au coin de l'ENSEEIH, ses relectures méticuleuses d'articles et de mémoire de thèse (quel courage !), nos randonnées pleines d'aventure (attention à la gadoue !), la piscine...

Mes remerciements vont aussi à Hélène pour tant d'années d'amitié et d'inoubliables moments passés ensemble.

Un grand merci à Marc qui a toujours répondu présent et su se rendre disponible quand il le fallait, pour son soutien tout au long de ces dernières années, pour ses réponses à mes innombrables questions.

Merci à toutes les personnes têtues et notamment au roi de ces personnes, qui se reconnaîtra, pour ne pas avoir lâché prise (d'accord, je viens d'écrire un pléonasme...).

Merci à Philippe pour ses relectures, pour m'avoir amenée en montagne et ainsi fait découvrir des lieux magiques et magnifiques (même par temps couvert !) dans les Pyrénées, pour m'avoir gentiment et régulièrement embêtée pendant cette thèse. Merci pour LA fameuse tarte aux myrtilles.

Merci à Némó (heu... Sandrine) pour tous les moments de rigolade qu'elle m'a offerts.

Un grand merci à Nassima pour tous les fous rires, ses invitations à « anger » et l'intrigue de la disparition de ses plantes, cet été, qui a égayé la rédaction de ce mémoire. Merci donc à toi « L'INTRUS. ».

Une pensée, à toi, Marius.

Merci à tous les membres de ma famille. Mes parents sans qui je n'en serais pas arrivée là, qui m'ont appris que rien ne s'acquiert facilement. Une pensée plus particulière à Maman, qui nous a quittés trop tôt et n'a malheureusement pas pu assister à cette fin de thèse. Merci aussi à mes grands-parents, à mes frères Christophe et Stéphane et à ma sœur Noëlle, à Catherine, à Aurélie et aux autres membres de la famille pour leur soutien et tout ce qui ne peut se nommer.

Pour finir, merci à toutes les personnes que j'ai malencontreusement oublié de remercier...

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Contexte . . . . .	1
1.2	Objectifs . . . . .	1
1.3	Liste des contributions identifiées . . . . .	3
1.4	Organisation du mémoire . . . . .	4
<b>2</b>	<b>Contexte : algorithme de suivi</b>	<b>5</b>
2.1	Cadre d'application . . . . .	5
2.1.1	Objet indéformable / déformable . . . . .	5
2.1.2	Scène indéformable / déformable . . . . .	6
2.1.3	Prise de vues . . . . .	7
2.2	La nature des connaissances et les hypothèses . . . . .	7
2.2.1	Les déformations . . . . .	7
2.2.2	La géométrie . . . . .	8
2.2.3	La photométrie . . . . .	8
2.3	Modélisation initiale . . . . .	8
2.3.1	Modèles de mouvement et paramètres . . . . .	8
2.3.1.1	Transformation rigide . . . . .	8
2.3.1.2	Transformation non rigide . . . . .	9
2.3.1.3	Exemple d'une image déformée par différents types de transformations . . . . .	14
2.3.2	Modèles de variations photométriques . . . . .	17
2.3.2.1	Invariance photométrique . . . . .	17
2.3.2.2	Modèle affine . . . . .	18
2.3.2.3	Modèles fondés sur une base d'apprentissage . . . . .	19
2.3.2.4	Ce que nous en retenons . . . . .	22
2.3.3	Occultations . . . . .	22
2.4	Suivi . . . . .	23
2.4.1	Notations . . . . .	23
2.4.2	Le problème fondamental ( $\mathcal{P}^s$ ) . . . . .	25
2.4.3	Quatre approches de résolution de ( $\mathcal{P}^s$ ) . . . . .	26
2.4.3.1	Résolution avec $\nabla I_t$ , sans composée . . . . .	26
2.4.3.2	Résolution sans $\nabla I_t$ , sans composée . . . . .	27
2.4.3.3	Résolution avec $\nabla I_t$ , avec composée . . . . .	29
2.4.3.4	Résolution sans $\nabla I_t$ , avec composée . . . . .	31
2.4.3.5	Bilan . . . . .	33
2.5	Suivi par prédiction linéaire . . . . .	34
2.5.1	Algorithme de Jurie et Dhome . . . . .	34

2.5.1.1	Principe informel . . . . .	34
2.5.1.2	Étape d'apprentissage . . . . .	34
2.5.1.3	Étape de suivi . . . . .	35
2.5.2	Notre variante (PP) . . . . .	35
2.5.2.1	Principe informel . . . . .	35
2.5.2.2	Étape d'apprentissage (PP_A) . . . . .	35
2.5.2.3	Étape de suivi (PP_S) . . . . .	36
2.5.2.4	Ce que nous retenons . . . . .	36
2.6	Bilan . . . . .	36
<b>3</b>	<b>Sélection de points d'intérêt</b>	<b>37</b>
3.1	Qu'est-ce qu'un point d'intérêt ? . . . . .	38
3.2	Les points d'intérêt et leurs détecteurs au fil du temps . . . . .	38
3.2.1	Détecteur de Beaudet . . . . .	39
3.2.1.1	La théorie . . . . .	39
3.2.1.2	La pratique . . . . .	40
3.2.2	Détecteur de Moravec . . . . .	43
3.2.2.1	La théorie . . . . .	43
3.2.2.2	La pratique . . . . .	45
3.2.2.3	Remarques . . . . .	45
3.2.3	Détecteur de Kitchen et Rosenfeld . . . . .	46
3.2.3.1	La théorie . . . . .	46
3.2.3.2	La pratique . . . . .	48
3.2.3.3	Remarques . . . . .	48
3.2.4	Détecteur de Harris . . . . .	49
3.2.4.1	La théorie . . . . .	49
3.2.4.2	La pratique . . . . .	52
3.2.4.3	Remarques . . . . .	54
3.2.5	Détecteur de Shi et Tomasi : KLT . . . . .	55
3.2.5.1	La théorie . . . . .	55
3.2.5.2	La pratique . . . . .	58
3.2.6	Détecteur de Smith et Brady : SUSAN . . . . .	59
3.2.6.1	La théorie . . . . .	59
3.2.6.2	La pratique . . . . .	62
3.2.7	Détecteur de Trajkovic et Hedley . . . . .	62
3.2.7.1	La théorie . . . . .	62
3.2.7.2	La pratique . . . . .	66
3.2.8	Détecteur de Lowe : SIFT . . . . .	67
3.2.8.1	La théorie . . . . .	67
3.2.8.2	La pratique . . . . .	68
3.2.8.3	Remarque . . . . .	70
3.2.9	Détecteur de Mikolajczyk et Schmid . . . . .	70
3.2.9.1	La théorie . . . . .	70
3.2.9.2	La pratique . . . . .	72
3.2.9.3	Comparaison avec SIFT . . . . .	72
3.2.10	Détecteur de Mikolajczyk et Schmid invariant aux transformations affines . . . . .	72
3.2.10.1	La théorie . . . . .	72
3.2.10.2	La pratique . . . . .	76

3.3	Détection pour suivre . . . . .	76
3.3.1	Shi et Tomasi : KLT . . . . .	76
3.3.2	Zivkovic et Van der Heijden . . . . .	76
3.3.2.1	Principe informel . . . . .	76
3.3.2.2	Algorithme (ZV) . . . . .	77
3.4	Sélection de sous-ensembles . . . . .	77
3.4.1	Sélection séquentielle par ajout : approche statistique . . . . .	77
3.4.1.1	Matrice de variances-covariances des paramètres d'un problème aux moindres carrés linéaires . . . . .	77
3.4.1.2	Une mesure de l'incertitude des paramètres estimés . . . . .	78
3.4.1.3	L'heuristique SFS . . . . .	78
3.4.1.4	Notre problème . . . . .	80
3.4.2	Sélection d'un sous-ensemble : approche algébrique . . . . .	83
3.4.2.1	Principe informel . . . . .	83
3.4.2.2	Algorithme de Golub <i>et al.</i> . . . . .	83
3.4.2.3	Algorithme de Ake Björck . . . . .	84
3.5	Sélection pour le suivi par prédiction linéaire . . . . .	84
3.5.1	La méthode de suivi . . . . .	84
3.5.2	Le problème . . . . .	85
3.5.3	L'algorithme général de sélection de points d'intérêt . . . . .	85
3.6	Implémentation des algorithmes de sélection des points d'intérêt . . . . .	85
3.6.1	Détermination de l'ensemble initial . . . . .	85
3.6.2	Sélection d'un sous-ensemble . . . . .	85
3.7	Expérimentation : la sélection de points d'intérêt et la validation des prédicteurs . . . . .	86
3.7.1	Les données . . . . .	86
3.7.2	Protocole . . . . .	86
3.7.2.1	Les ensembles initiaux . . . . .	86
3.7.2.2	Les filtres . . . . .	88
3.7.2.3	Les déformations . . . . .	90
3.7.2.4	Les séquences artificielles . . . . .	91
3.7.2.5	Validation . . . . .	93
3.7.3	Résultats . . . . .	94
3.7.3.1	Concernant la validation du prédicteur . . . . .	94
3.7.3.2	Concernant le suivi de séquences artificielles . . . . .	102
3.8	Bilan . . . . .	103
<b>4</b>	<b>Suivi d'objets déformables</b> . . . . .	<b>105</b>
4.1	Généralités sur les maillages . . . . .	105
4.1.1	Définitions relatives aux maillages . . . . .	105
4.1.2	Topologie et géométrie d'un maillage . . . . .	111
4.1.2.1	Choix de la géométrie d'un maillage . . . . .	111
4.1.2.2	Construction d'un maillage plan 2D . . . . .	113
4.1.2.3	Modifications de la géométrie ou de la topologie d'un maillage . . . . .	116
4.1.3	Interpolation . . . . .	119
4.1.3.1	Définitions et propriétés . . . . .	120
4.1.3.2	Application de l'interpolation au maillage . . . . .	121
4.1.4	Les coordonnées barycentriques . . . . .	123

4.1.4.1	Choix de l'espace affine . . . . .	123
4.1.4.2	Choix de la base affine . . . . .	123
4.1.4.3	Définition des coordonnées barycentriques canoniques . .	123
4.1.4.4	Correspondance entre coordonnées cartésiennes et coordonnées barycentriques canoniques . . . . .	124
4.1.4.5	Définition des coordonnées barycentriques généralisées .	124
4.1.4.6	Utilisation des coordonnées barycentriques généralisées .	126
4.1.5	Transformations géométriques d'un maillage, utiles au suivi d'objets	126
4.1.5.1	Choix des paramètres . . . . .	126
4.1.5.2	Transformations globales . . . . .	127
4.1.5.3	Transformations nodales / nodales inverses . . . . .	128
4.1.5.4	Application d'une transformation nodale inverse à un maillage dépendant de l'original . . . . .	129
4.1.5.5	Compositions exacte et approchée d'une transformation nodale et d'une transformation globale définies sur le même maillage . . . . .	133
4.2	Maillage et suivi : état de l'art . . . . .	137
4.2.1	Méthodes mono-problème (globales) . . . . .	137
4.2.2	Méthodes pluri-problèmes (locales) . . . . .	138
4.2.2.1	Algorithme de Nakaya et Harashima . . . . .	138
4.2.2.2	Algorithmes d'Altunbasak et Tekalp . . . . .	139
4.3	Maillage et suivi d'objets déformables : notre proposition . . . . .	144
4.3.1	Introduction . . . . .	144
4.3.1.1	Nos objectifs . . . . .	144
4.3.1.2	Nos hypothèses . . . . .	144
4.3.1.3	Nos leviers . . . . .	145
4.3.2	À propos de l'apprentissage artificiel . . . . .	145
4.3.2.1	La théorie . . . . .	145
4.3.2.2	La pratique . . . . .	146
4.3.3	À propos du suivi « au fil de l'eau » . . . . .	146
4.3.3.1	La théorie . . . . .	146
4.3.3.2	La pratique . . . . .	148
4.3.4	À propos de la complexité . . . . .	150
4.4	Expérimentation . . . . .	151
4.4.1	Protocole . . . . .	151
4.4.1.1	Les caractéristiques communes . . . . .	152
4.4.1.2	Les séquences d'images . . . . .	153
4.4.1.3	Les algorithmes de suivi comparés . . . . .	154
4.4.1.4	Validation . . . . .	155
4.4.2	Résultats . . . . .	156
4.4.2.1	Validation locale sur des séquences d'images artificielles	156
4.4.2.2	Validation globale sur des séquences d'images artificielles	159
4.4.2.3	Validation globale sur des séquences d'images réelles . .	162
4.5	Bilan . . . . .	165

<b>5 Conclusion et perspectives</b>	<b>167</b>
5.1 Bilan . . . . .	167
5.2 Contributions . . . . .	168
5.3 Perspectives . . . . .	169
5.4 Les articles . . . . .	169
<b>A Rappel sur les polynômes orthogonaux discrets de Tchebychev : la théorie</b>	<b>171</b>
A.1 Cas de la dimension 1 . . . . .	171
A.1.1 Une base de polynômes orthogonaux discrets de Tchebychev . . . . .	171
A.1.2 Le polynôme interpolateur de f . . . . .	172
A.2 Cas de la dimension 2 . . . . .	172
A.3 Exemples . . . . .	173
<b>B Le mythe de l'opérateur « SOBEL »</b>	<b>175</b>
B.1 La genèse d'un opérateur mythique . . . . .	175
B.2 Les fondements mathématiques de l'opérateur « SOBEL » . . . . .	176
<b>C Courbure d'une surface</b>	<b>179</b>
C.1 Les divers composants de base . . . . .	179
C.2 L'endomorphisme de Weingarten . . . . .	181
C.2.1 Définition . . . . .	181
C.2.2 Propriétés . . . . .	181
C.3 Directions principales . . . . .	181
C.4 Courbures moyenne et gaussienne . . . . .	182
C.4.1 Les composants introduits par Gauss . . . . .	182
C.4.2 Courbure moyenne : $\kappa_m(M)$ . . . . .	182
C.4.3 Courbure gaussienne : $\kappa_g(M)$ . . . . .	182
C.5 Courbures normale et principales . . . . .	183
C.5.1 Courbure normale . . . . .	183
C.5.2 Courbures principales . . . . .	183





# Chapitre 1

## Introduction

### 1.1 Contexte

En traitement d'images et en vision par ordinateur, le suivi d'objets en mouvement dans un environnement connu ou inconnu est couramment étudié. Il s'appuie sur des propriétés d'invariance des objets d'intérêt. L'invariance peut concerner la géométrie de la scène ou des objets, l'apparence des objets (c'est-à-dire la photométrie ou la couleur) ou encore la cinématique (par exemple les contraintes spatio-temporelles).

En vision artificielle, le suivi d'objets est au cœur de nombreuses applications. Donnons deux exemples. Le suivi peut être un outil pour donner des capacités d'autonomie visuelle à des automates ou des robots. Dans ce cas, la perception visuelle est un préalable à l'action et nécessite un apprentissage pour établir des liens entre les causes et les actions à produire en réponse. Le suivi d'objets peut aussi permettre d'automatiser des tâches répétitives. Par exemple, dans le cas de la vidéo-surveillance, la perception visuelle a pour objectif d'aider l'être humain et de l'informer d'un problème en détectant des événements anormaux. Cela nécessite donc de nouveau un apprentissage. Dans ces deux exemples, l'apprentissage se situe essentiellement au niveau de l'interprétation du suivi. De la même manière, l'apprentissage peut intervenir dans l'algorithme de suivi afin de prédire le mouvement des objets. Dans ce cas, l'apprentissage s'effectue sur un préfixe d'une séquence d'images afin de réaliser le suivi sur la séquence complète ou sur une autre séquence.

Fondamentalement, le suivi n'est que l'appariement de zones d'images ou l'appariement de points d'intérêt. L'appariement est un domaine étudié depuis des décennies en vision par ordinateur et qui reste fort actif de nos jours tant les problèmes sous-jacents sont compliqués.

### 1.2 Objectifs

À partir d'une séquence d'images représentant un objet déformable en mouvement, nous cherchons à déterminer, le plus efficacement et précisément possible, sa position dans chaque image et donc à déterminer son déplacement ou encore son mouvement apparent dans l'image ou sa déformation d'une image à l'autre.

Pour cela, nous cherchons tout d'abord à déterminer quelles sont les « bonnes » représentations visuelles des objets en mouvement filmés par une caméra monoculaire. Nous faisons l'hypothèse que la caméra, aussi bien que les objets, peut bouger et que les mouvements apparents sont de petite amplitude. Cette dernière hypothèse n'est pas restrictive.

En effet, en filmant avec une prise d'images plus fréquente, les amplitudes du mouvement apparent deviennent plus faibles.

L'objet suivi est initialement un objet tridimensionnel. La modélisation peut donc s'opérer au niveau tridimensionnel ou bidimensionnel. Dans ce dernier cas, l'étude porte sur la représentation de la projection de l'objet dans le plan image.

Dans notre étude, nous ne nous intéressons pas à la représentation 3D de l'objet, et donc à déterminer ou utiliser un modèle tridimensionnel de cet objet (comme par exemple un modèle CAO, un maillage 3D, etc.), mais plutôt à une représentation 2D de l'objet 3D projeté dans le plan image.

Cette représentation bidimensionnelle peut être définie à partir d'un ensemble de caractéristiques visuelles denses ou discrètes. Une représentation dense possède l'avantage de fournir en détail toutes les informations relatives à l'objet mais a l'inconvénient d'être lourde à manipuler, notamment dans le cas d'un suivi rapide. Dans le cas d'une représentation discrète, les caractéristiques peuvent être structurées ou non et ceci spatialement ou temporellement. Par exemple, une structure spatiale de maillage peut être utilisée.

Le dernier point à étudier est la prise en compte des variations de forme, de mouvement ou d'apparence des objets dans la représentation. Un objet tridimensionnel peut être de forme fixe ou au contraire déformable. Il peut se déplacer selon un mouvement contraint ou arbitraire. L'intérêt peut aussi se limiter au mouvement projeté sur le plan image sans autre hypothèse que la continuité de celui-ci. Les autres caractéristiques (couleur, luminosité, transparence...) sont regroupées sous le terme d'apparence. Pour nous, l'invariance de l'apparence concerne l'invariance des niveaux de gris. Comme cela est illustré sur la figure 1.1, les niveaux de gris sous les points correspondants sont les mêmes.

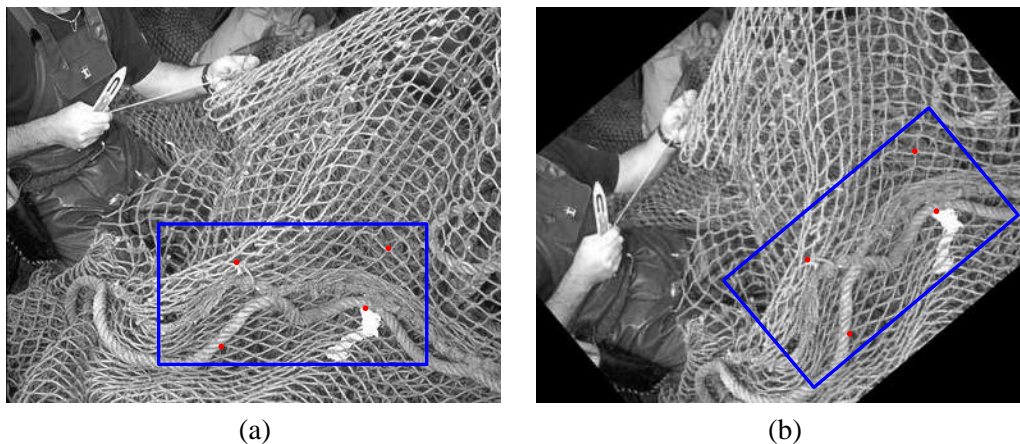


FIG. 1.1 – Le rectangle bleu représente la zone d'intérêt, les points rouges représentent les points d'intérêt. (a) L'image initiale, (b) l'image déformée. L'apparence de la zone d'intérêt est constante.

Nous défendons l'utilisation d'une représentation discrète, dans l'espace image, des objets en mouvement arbitraire, sous l'hypothèse de l'invariance des niveaux de gris, par :

- des points d'intérêt, caractéristiques de la texture de l'image, obtenus par un détecteur de points. Ce choix est guidé par les propriétés d'invariance aux transformations affines, aux changements d'échelle, etc. des points d'intérêt. Ces propriétés dépendent évidemment du détecteur de points utilisé ;

- des maillages bidimensionnels de la projection de l'objet 3D sur le plan image. Ce choix est guidé par la possibilité de pouvoir passer rapidement d'une représentation discrète, la représentation d'un maillage défini par ses nœuds, à une représentation continue via une interpolation sur tout le maillage. Cet outil possède aussi l'avantage de conserver, tout au long du suivi, une unité dans l'objet. L'objet est ainsi représenté de manière compacte. Ce modèle est également d'intérêt en analyse et en synthèse d'images. A priori, les nœuds du maillage n'ont aucune raison d'être des points caractéristiques de la texture, ni d'être les projections d'un hypothétique modèle 3D de l'objet.

Notre premier axe de recherche a pour objectif d'améliorer l'algorithme de suivi d'objets plans indéformables de Jurie et Dhome [JD02], référencé (JD) dans ce mémoire de thèse, en sélectionnant méticuleusement les caractéristiques visuelles sur lesquelles il se fonde. Cet algorithme utilise à bon escient l'hypothèse de l'invariance de l'apparence ou encore de l'invariance des niveaux de gris en un certain nombre de points d'intérêt et part du principe qu'il y a équivalence entre « déformer une image et s'intéresser à un ensemble de points d'intérêt immobiles » et « déformer un ensemble de points d'intérêt et s'intéresser à une image constante ». La figure 1.1 montre deux images déformées avec leur ensemble de points d'intérêt et met ainsi en évidence que l'apparence de la zone d'intérêt sur les deux images est la même. Ainsi, tout changement entre deux images est source d'information fort utile pour prédire le mouvement réel. Pour cela, nous étudions un processus de sélection de points d'intérêt adaptés à cet algorithme.

Notre second axe de recherche concerne le suivi d'objets déformables en utilisant, pour notre algorithme, le principe de l'algorithme de suivi (JD) et un maillage pour représenter les objets. Le choix de la représentation par maillage vient du fait qu'un maillage possède un grand nombre de degrés de liberté et donc la capacité d'approcher au mieux le mouvement apparent de l'objet. Il permet aussi de représenter l'objet et son mouvement sur peu de données tout en conservant la structure de l'objet. L'algorithme de suivi (JD) est, quant à lui, adapté au cas du suivi par transformation d'un maillage. Il sera ainsi appliqué à chaque nœud du maillage afin d'en conserver la topologie.

### 1.3 Liste des contributions identifiées

Tous les travaux de ce mémoire de thèse sont donc à raccorder directement ou indirectement à l'algorithme de suivi (JD) et à la sélection de points d'intérêt.

L'algorithme de suivi (JD) est constitué d'une étape d'apprentissage hors ligne, qui estime un prédicteur linéaire, et d'une étape de suivi en ligne. Nous en améliorons l'efficacité en reportant une partie des calculs de l'étape de suivi vers l'étape d'apprentissage et l'exactitude en sélectionnant soigneusement l'ensemble des points déterminant le prédicteur linéaire. Nous identifions trois contributions :

- 1) Une variante de l'algorithme de suivi (JD). Nous la nommons (PP). L'objectif est de diminuer le temps de calcul lors de l'étape de suivi (voir §2.5.2). Cette variante est de même fortement utile lorsque les transformations du modèle de mouvement ne forment pas un groupe (par exemple, les transformations d'un maillage). Cette approche permet ainsi de limiter le nombre d'approximations.
- 2) L'utilisation originale, dans le domaine du traitement d'images et de la vision, de méthodes statistique et algébrique de sélection d'un sous-ensemble de données, que nous appliquons à la sélection d'un sous-ensemble de points d'intérêt (voir §3.4).

- 3) Un nouvel algorithme de sélection d'un sous-ensemble de points d'intérêt fondé sur une sélection séquentielle par ajout des points d'intérêt (voir §3.4.1.4). Ceci permet d'améliorer la précision du suivi.

Concernant le suivi d'objets déformables en utilisant un maillage et en adaptant l'algorithme de suivi (JD) dans le cas de transformations ne formant pas un groupe, nous identifions trois contributions :

- 4) La caractérisation géométrique et analytique de la transformation élémentaire d'un maillage que nous appelons transformation nodale (voir §4.1.5.3) ainsi que celle de son inverse (voir §4.1.5.4). Elle est utilisée lors de l'étape d'apprentissage.
- 5) Une nouvelle approximation de la composition d'une transformation globale et d'une transformation nodale d'un maillage en mettant en œuvre les coordonnées barycentriques généralisées (voir §4.1.5.5). Elle est utilisée lors de l'étape de suivi.
- 6) Un nouvel algorithme de suivi d'objets déformables utilisant un maillage et fondé sur notre variante (PP) de l'algorithme de suivi (JD), ainsi qu'un schéma itératif pour prendre en compte tous les nœuds. Nous le nommons (CMPT) (voir §4.3). Il est plus rapide que les autres algorithmes de suivi étudiés utilisant un maillage.

## 1.4 Organisation du mémoire

Ce mémoire de thèse s'organise comme suit :

- 1) Le chapitre 2 présente des méthodes de suivi (contribution 1). Pour cela, les définitions des termes couramment utilisés dans ce mémoire et les éléments de base utiles pour les méthodes de suivi sont donnés, de même que des méthodes de suivi avec en particulier la méthode de suivi (JD) et notre variante (PP) ;
- 2) Le chapitre 3 traite de l'amélioration de la méthode de suivi (JD) par sélection de points d'intérêt (contributions 2 et 3). Ainsi, sont présentés un historique des détecteurs de points d'intérêt de 1976 à nos jours, des filtres permettant de cribler les points d'intérêt et des résultats comparatifs ;
- 3) Le chapitre 4 est axé sur le suivi d'objets déformables en utilisant un maillage (contributions 4, 5 et 6). Son contenu intègre en particulier une description des maillages et de leurs transformations, restreintes à celles que nous utilisons, une présentation de notre algorithme de suivi (CMPT) et des résultats comparatifs avec d'autres méthodes ;
- 4) Le chapitre 5 présente une conclusion de ces études et évoque des perspectives.

## Chapitre 2

# Contexte : algorithme de suivi

Ce chapitre traite d'algorithmes de suivi visuel, qu'il s'agisse du suivi de :

- points indépendamment les uns des autres,
- régions,
- ensembles de points.

Certaines procédures de suivi sont adaptées au suivi d'**objets indéformables** ou **déformables**. Enfin, dans le cas où le mouvement est modélisé par une transformation paramétrique **rigide** ou **non rigide**, le suivi consiste à estimer ces paramètres.

Ce chapitre s'attelle à présenter, à partir du problème fondamental de suivi, différentes méthodes pour le résoudre ainsi qu'une méthode de suivi faisant intervenir un prédicteur linéaire.

Notre contribution est la mise en œuvre d'une variante du suivi par prédiction linéaire.

Ce chapitre s'articule comme suit :

- une description des objets à suivre et des hypothèses de suivi faites ;
- une présentation des modèles de mouvement et leurs paramètres ;
- une description de quatre approches de résolution du problème fondamental de suivi ;
- une description d'un algorithme de suivi par prédiction linéaire et de notre variante.

## 2.1 Cadre d'application

### 2.1.1 Objet indéformable / déformable

Dans le dictionnaire « Le Petit Robert » [Rob03], un objet est défini comme étant une « chose solide ayant unité et indépendance et répondant à une certaine destination ».

D'après cette définition, les liquides, les fumées ou gaz, le sable, etc. ainsi que toute forme composée de plusieurs choses non définitivement liées entre elles ne sont pas considérés comme des objets.

Dans la réalité, tous les objets 3D sont plus ou moins déformables. Cependant, lorsque les propriétés physiques d'un objet contraignent ses déformations internes à être d'amplitudes très faibles, il est qualifié d'**indéformable** (voir définition 2.1.1.1). C'est le cas, par exemple, d'une statue en fer qui, sous l'effet de la chaleur, se dilate.

### 2.1.1.1 Définition (Objet indéformable)

Un objet  $\mathcal{O}$  est dit **indéformable** si la distance entre deux quelconques de ses points est constante au cours du temps.

$$\forall P, Q \in \mathcal{O}, P \neq Q, \forall t > 0, \|\overrightarrow{PQ}(t)\| = \|\overrightarrow{PQ}(0)\| = cste. \quad (2.1)$$

### 2.1.1.2 Définition (Objet déformable)

Un objet  $\mathcal{O}$  est dit **déformable** s'il existe au moins deux de ses points dont la distance varie au cours du temps.

$$\exists P, Q \in \mathcal{O}, P \neq Q, \exists t, \exists t' \neq t \text{ tels que } \|\overrightarrow{PQ}(t)\| \neq \|\overrightarrow{PQ}(t')\|. \quad (2.2)$$

Dans son mémoire de thèse, Julien Lenoir [Len04] distingue deux types d'objets déformables : ceux qui possèdent une structure interne (par exemple : un cube en bois) et ceux qui n'en ont pas ; ils sont alors représentés par des particules (par exemple : le sable, la fumée...).

Dans ce mémoire de thèse, nous ne nous intéressons qu'aux objets possédant une structure interne.

Il est à noter que certains objets peuvent être qualifiés de déformable ou non en fonction de la tolérance accordée pour passer de la caractérisation mathématique (équation (2.1)) à la mesure physique. Dans notre cas, nous considérons qu'une aile d'avion est un objet indéformable alors qu'un ingénieur étudiant la résistance de celle-ci en vol la considère comme déformable.

De la même façon, un objet qualifié de déformable peut être momentanément considéré comme indéformable, si, pendant le temps de l'étude, il n'est pas déformé. Nous pouvons donner l'exemple d'une voiture qui est considérée comme indéformable si les portières et les vitres ne sont pas actionnées pendant le temps de l'étude et comme déformable sinon.

Dresser une liste exhaustive d'objets indéformables ou, par élimination, une liste d'objets déformables n'est pas possible ; voici tout de même quelques exemples :

- objets indéformables : statue, voiture, maison, pierre, bois...
- objets déformables : personne, feuille de papier, tissus, ballon de baudruche...

## 2.1.2 Scène indéformable / déformable

Tout comme un objet, une scène peut être qualifiée d'indéformable ou de déformable.

### 2.1.2.1 Définition (Scène indéformable)

Une **scène indéformable** est uniquement composée d'**objets indéformables** rigidement liés entre eux.

### 2.1.2.2 Définition (Scène déformable)

Une **scène déformable** peut provenir de différents assemblages d'objets indéformables ou déformables, en mouvement ou non.

Une scène déformable peut être composée :

- uniquement d'**objets indéformables** en mouvement les uns par rapport aux autres. Par exemple : le trafic routier à l'entrée d'un tunnel...
- uniquement d'**objets déformables** éventuellement en mouvement. Par exemple : un groupe de personnes en train de parler, un essaim d'abeilles à la recherche d'un endroit où se poser...
- d'un mélange d'**objets indéformables** et **déformables** éventuellement en mouvement. Par exemple : un oiseau (objet déformable) se balançant (mouvement) sur un trapèze (objet indéformable), une vache (objet déformable) dans un pré à côté d'un train (objet indéformable) qui passe (mouvement), des enfants (objets déformables) dans des autos tamponneuses (objets indéformables) en mouvement...

Par exemple, dans leurs travaux [JSR05], Jacquot *et al.* ont étudié le suivi au sein d'une scène déformable d'objets indéformables (une voiture [[JSR05], fig. 4 et 5, p. 108]) et d'objets déformables (le visage d'une personne [[JSR05], fig. 3, p. 107] ou un joueur de football [[JSR05], fig. 6, p. 109]).

### 2.1.3 Prise de vues

Pour passer de l'espace tridimensionnel à une de ses représentations sous forme d'image, un système de prise de vues est utilisé. Dans notre cas, il s'agit d'une caméra dite en trou d'épingle. Le modèle mathématique sous-jacent est une projection centrale sur un plan. Dans cette étude, nous nous intéressons aux transformations bidimensionnelles entre deux images d'une séquence d'images. Ces transformations sont induites :

- par le mouvement dans la scène,
- ou par le mouvement de la caméra.

La transformation entre les deux images est généralement non rigide (cf. définition 2.3.1.2.1) sauf :

- pour le cas très particulier où la caméra est immobile et l'objet subit une rotation autour de l'axe principal (c'est-à-dire l'axe passant par le centre optique et le point principal) ;
- pour le cas où la scène et la caméra subissent un mouvement de translation commun.

## 2.2 La nature des connaissances et les hypothèses

Des hypothèses sur l'évolution de la cible 2D (un objet par exemple) sont indispensables pour rendre possible le suivi.

### 2.2.1 Les déformations

L'aspect de la cible est supposé ne pas changer du tout au tout d'une image à l'autre. L'amplitude des déplacements des points de la cible est donc faible et bornée. L'hypothèse de petits déplacements est faite.

### 2.2.2 La géométrie

La géométrie décrit les déplacements de chacun des points de la cible que l'on suit au fil des images d'une séquence d'images. Ces déplacements sont souvent limités en amplitude.

Habituellement, une transformation, appelée mouvement, modélise les déplacements de tous les points de la région cible. Cette fonction notée  $T$  dépend d'un vecteur de paramètres  $\beta_t \in \mathbb{R}^p$  qui évolue au cours de la séquence. Le modèle de mouvement s'écrit :

$$\mathbf{m}' = T(\mathbf{m}, \beta_t).$$

### 2.2.3 La photométrie

La photométrie concerne l'aspect de la cible. Le plus souvent, elle stipule que les niveaux de gris de la cible ne changent pas entre deux images. Cependant d'autres hypothèses peuvent être effectuées (voir §2.3.2).

## 2.3 Modélisation initiale

### 2.3.1 Modèles de mouvement et paramètres

Dans les méthodes de suivi, les modèles de mouvement sont représentés par des transformations paramétrées. Cette section contient :

- la définition des transformations rigides ;
- la définition des transformations non rigides ;
- des exemples de telles transformations.

#### 2.3.1.1 Transformation rigide

Le mouvement d'un objet 3D peut être mathématiquement modélisé par une transformation.

##### 2.3.1.1.1 Définition (Transformation rigide)

*Une **transformation rigide** est une transformation qui conserve à la fois les distances (c'est une isométrie) et l'orientation (l'image d'un repère orthonormé direct est un repère orthonormé direct).*

Les transformations rigides élémentaires sont :

- la transformation identité,
- la rotation,
- et la translation.

Une **transformation rigide** est donc le résultat de la composition de transformations élémentaires. Il est à noter que, dans l'espace tridimensionnel, un objet indéformable ne peut être déplacé que selon une transformation rigide. Son **mouvement** est qualifié de rigide.



### 2.3.1.2 Transformation non rigide

#### 2.3.1.2.1 Définition (Transformation non rigide)

Est **non rigide** une transformation qui ne conserve pas les distances ou (non exclusif) les orientations.

C'est le cas de la majorité des transformations qui permettent de modéliser des déformations physiques.

Les modèles mathématiques sont d'ordre supérieur du type polynôme, spline, combinaison linéaire de fonctions de base...

Dans ce mémoire de thèse, après avoir présenté quelques-unes de ces transformations, nous faisons essentiellement référence aux transformations affine et homographique, qui sont largement utilisées dans notre domaine de recherche.

Suit une liste non exhaustive de transformations non rigides :

- **transformation affine** : elle conserve le parallélisme. Ce type de transformation est utilisé dans le cadre du suivi par [HB98, LK81, AT97b].

Dans l'espace bidimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^2 \times \mathbb{R}^6 \rightarrow \mathbb{R}^2 \\ (p = \begin{bmatrix} x \\ y \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \beta_1 x + \beta_2 y + \beta_3 \\ \beta_4 x + \beta_5 y + \beta_6 \end{bmatrix} \\ \\ = \underbrace{\begin{bmatrix} \beta_1 & \beta_2 \\ \beta_4 & \beta_5 \end{bmatrix}}_{A_{2d}} p + \underbrace{\begin{bmatrix} \beta_3 \\ \beta_6 \end{bmatrix}}_{t_{2d}} \end{array} \right. \quad (2.3)$$

Dans l'espace tridimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^3 \times \mathbb{R}^{12} \rightarrow \mathbb{R}^3 \\ (p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \beta_1 x + \beta_2 y + \beta_3 z + \beta_4 \\ \beta_5 x + \beta_6 y + \beta_7 z + \beta_8 \\ \beta_9 x + \beta_{10} y + \beta_{11} z + \beta_{12} \end{bmatrix} \\ \\ = \underbrace{\begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \\ \beta_5 & \beta_6 & \beta_7 \\ \beta_9 & \beta_{10} & \beta_{11} \end{bmatrix}}_{A_{3d}} p + \underbrace{\begin{bmatrix} \beta_4 \\ \beta_8 \\ \beta_{12} \end{bmatrix}}_{t_{3d}} \end{array} \right. \quad (2.4)$$

- **transformation homographique** : elle conserve l'alignement des points. Ce type de transformation est utilisé pour modéliser le passage de l'espace tridimensionnel à l'espace bidimensionnel et dans le cadre du suivi par [HB98, JD02, Dou04, RH95].

Dans l'espace bidimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^2 \times \mathbb{R}^8 \rightarrow \mathbb{R}^2 \\ (p = \begin{bmatrix} x \\ y \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \beta_1 x + \beta_2 y + \beta_3 \\ \beta_7 x + \beta_8 y + 1 \\ \beta_4 x + \beta_5 y + \beta_6 \\ \beta_7 x + \beta_8 y + 1 \end{bmatrix} \end{array} \right. \quad (2.5)$$

Dans l'espace tridimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^3 \times \mathbb{R}^{15} \rightarrow \mathbb{R}^3 \\ (p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \frac{\beta_1 x + \beta_2 y + \beta_3 z + \beta_4}{\beta_{13} x + \beta_{14} y + \beta_{15} z + 1} \\ \frac{\beta_5 x + \beta_6 y + \beta_7 z + \beta_8}{\beta_{13} x + \beta_{14} y + \beta_{15} z + 1} \\ \frac{\beta_9 x + \beta_{10} y + \beta_{11} z + \beta_{12}}{\beta_{13} x + \beta_{14} y + \beta_{15} z + 1} \end{bmatrix} \end{array} \right. \quad (2.6)$$

De l'espace tridimensionnel à l'espace bidimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^3 \times \mathbb{R}^{11} \rightarrow \mathbb{R}^2 \\ (p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \frac{\beta_1 x + \beta_2 y + \beta_3 z + \beta_4}{\beta_9 x + \beta_{10} y + \beta_{11} z + 1} \\ \frac{\beta_5 x + \beta_6 y + \beta_7 z + \beta_8}{\beta_9 x + \beta_{10} y + \beta_{11} z + 1} \end{bmatrix} \end{array} \right. \quad (2.7)$$

- **transformation quadratique ;**

Dans l'espace bidimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^2 \times \mathbb{R}^{12} \rightarrow \mathbb{R}^2 \\ (p = \begin{bmatrix} x \\ y \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \beta_1 x^2 + \beta_2 y^2 + \beta_3 xy + \beta_4 x + \beta_5 y + \beta_6 \\ \beta_7 x^2 + \beta_8 y^2 + \beta_9 xy + \beta_{10} x + \beta_{11} y + \beta_{12} \end{bmatrix} \end{array} \right. \quad (2.8)$$

Dans l'espace tridimensionnel :

$$\left\{ \begin{array}{l} T : \mathbb{R}^3 \times \mathbb{R}^{30} \rightarrow \mathbb{R}^3 \\ (p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \beta) \mapsto T(p, \beta) = \begin{bmatrix} \beta_1 x^2 + \beta_2 y^2 + \beta_3 z^2 + \beta_4 xy + \beta_5 xz + \beta_6 yz + \beta_7 x + \beta_8 y + \beta_9 z + \beta_{10} \\ \beta_{11} x^2 + \beta_{12} y^2 + \beta_{13} z^2 + \beta_{14} xy + \beta_{15} xz + \beta_{16} yz + \beta_{17} x + \beta_{18} y + \beta_{19} z + \beta_{20} \\ \beta_{21} x^2 + \beta_{22} y^2 + \beta_{23} z^2 + \beta_{24} xy + \beta_{25} xz + \beta_{26} yz + \beta_{27} x + \beta_{28} y + \beta_{29} z + \beta_{30} \end{bmatrix} \end{array} \right. \quad (2.9)$$

- **transformation polynomiale (de degré > 2) ;**

- **transformation fondée sur des fonctions de base radiales :** Une fonction de base  $\phi$ , radiale non linéaire, induit le type de transformation non rigide. Ces transformations sont très utilisées dans le cadre de l'interpolation de surface. Une fonction radiale est une fonction dont la valeur ne dépend que de la distance à l'origine ou à un autre centre d'influence  $c$ . Elle est donc évaluée respectivement en  $\|x\|$  ou en  $\|x - c\|$ . Généralement, la norme utilisée est la norme euclidienne. Voici une liste non exhaustive de fonctions de base radiales les plus couramment utilisées,  $\phi$  peut être :

\* une gaussienne :

$$\left\{ \begin{array}{l} \phi : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right). \end{array} \right. \quad (2.10)$$

La figure 2.1-(a) représente l'application «  $p \in \mathbb{R}^2 \mapsto \phi(\|p\|)$  », où l'écart type  $\sigma$  vaut 1. Ce type de fonction a donc un impact local autour du centre d'influence

(ici,  $c = [0\ 0]^T$ ). Ces fonctions gaussiennes sont les plus utilisées car elles sont les plus vraisemblables physiquement et ont une réponse finie ce qui est plus pratique.

- \* une spline de type « plaque mince » (TPS : *Thin Plate Spline*) [[Boo89], §I-D, p. 570]. Ces fonctions sont dues à Duchon [Duc76] :

$$\begin{cases} \phi : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \phi(x) = \begin{cases} x^2 \ln(x) & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \end{cases} \quad (2.11)$$

La figure 2.1-(b) représente l'application «  $p \in \mathbb{R}^2 \mapsto -\phi(\|p\|)$  ». Plus la distance du point au centre (ici  $[0\ 0]^T$ ) est grande, plus la valeur (absolue) et donc l'impact de cette fonction sont grands. Ce type de fonction a donc une influence globale.

- \* une fonction quadratique :

$$\begin{cases} \phi : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \phi(x) = \sqrt{x^2 + a^2}, \end{cases} \quad (2.12)$$

où  $a \in \mathbb{R}$ . Hardy [Har71] a introduit en 1971 le terme « multiquadratique » pour désigner une combinaison linéaire [[Har71], éq. (2), p. 1906] de telles fonctions quadratiques. Il les a utilisées pour définir un modèle analytique de terrain.

La figure 2.1-(c) représente l'application «  $p \in \mathbb{R}^2 \mapsto -\phi(\|p\|)$  », où  $\phi$  est définie par l'équation (2.12). Plus la distance du point au centre (ici  $[0\ 0]^T$ ) est grande, plus la valeur et donc l'impact de cette fonction sont grands. Ce type de fonction a donc également une influence globale.

La figure 2.1 illustre une combinaison linéaire de telles fonctions de base (figure 2.1-(d) pour les fonctions de base de type gaussienne, 2.1-(e) pour celles de type « plaque mince », et 2.1-(f) pour celles de type quadratique) :

$$\begin{cases} g : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto g(x, y) = \phi(\|x - c_1\|) - \phi(\|x - c_2\|) \\ \quad + \phi(\|x - c_3\|) - \phi(\|x - c_4\|), \end{cases} \quad (2.13)$$

où les  $\{c_i\}_{i=1..4}$  sont les centres d'influence, avec :

$$\begin{cases} c_1 = [0\ 1]^T \\ c_2 = [1\ 0]^T \\ c_3 = [0\ -1]^T \\ c_4 = [-1\ 0]^T. \end{cases}$$

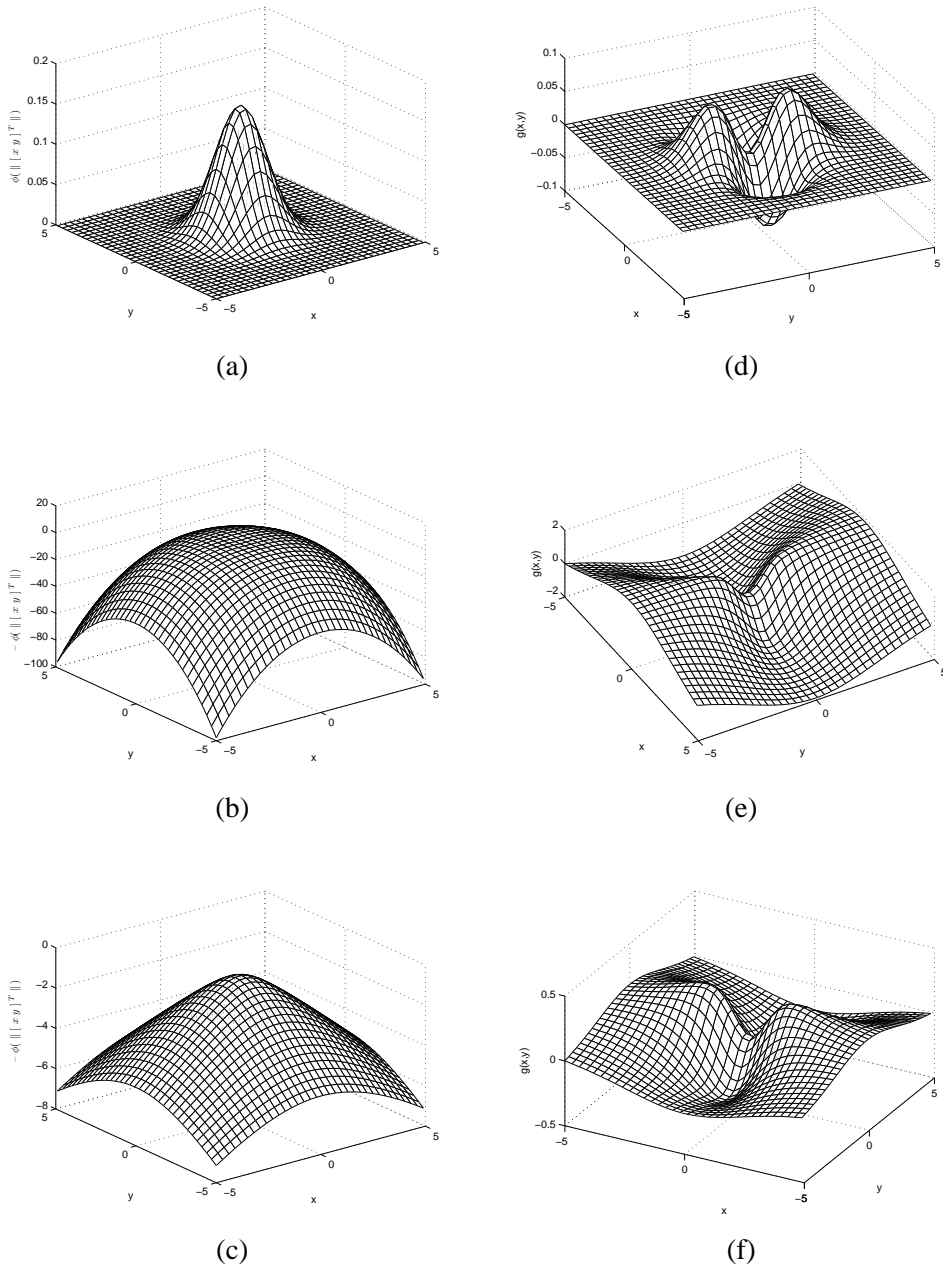


FIG. 2.1 – Colonne 1 : Fonctions de base : (a) gaussienne ; (b) spline de type « plaque mince » (TPS) ; (c) quadratique. Colonne 2 : Exemples de transformations fondées sur des fonctions de base radiales : (d) fonction multigaussienne ; (e) fonction multiplaqueminces ; (f) fonction multiquadratique

Au sens de Hardy,  $g$  est donc respectivement une fonction « multigaussienne », « multiplaqueminces », « multiquadratique ». Pour les fonctions de type gaussienne (figure 2.1-(d)), l'écart type  $\sigma$  est fixé à 1. Pour les fonctions de type quadratique, le paramètre  $a$  est fixé à 100. Il est à noter que plus les centres sont éloignés, plus l'influence des uns sur les autres est forte dans le cas des splines de type « plaque mince » ou quadratique. Ce n'est pas vrai pour les gaussiennes.

Cette transformation fondée sur des fonctions de base radiales est la solution d'un

problème d'approximation d'une surface dont la connaissance de certains points  $c_k$  de la surface est donnée :  $\forall k, T(c_k, \beta) = b_{c_k}$ . La surface doit impérativement passer par ces points. Dans le cas des fonctions gaussiennes (équation (2.10)), les coefficients  $w_k$  de la transformation :

$$\begin{cases} T : \mathbb{R}^2 \times \mathbb{R}^{2l+2l} & \rightarrow \mathbb{R}^2 \\ (p, \beta) & \mapsto T(p, \beta) = \sum_{k=1}^l w_k \phi(\|p - c_k\|) \end{cases} \quad (2.14)$$

peuvent être déterminés de manière unique [Buh00]. Dans le cas des fonctions de type spline de type « plaque mince » (équation (2.11)) et quadratiques (équation (2.12)), la solution n'est pas unique sauf s'il est rajouté un polynôme  $C$  d'ordre  $d$  correctement choisi [Buh00]. La transformation devient donc :

$$\begin{cases} T : \mathbb{R}^2 \times \mathbb{R}^{(d+1)*(d+2)+2l+2l} & \rightarrow \mathbb{R}^2 \\ (p, \beta) & \mapsto T(p, \beta) = C(p) + \sum_{k=1}^l w_k \phi(\|p - c_k\|). \end{cases} \quad (2.15)$$

Dans le cas des fonctions TPS, un polynôme d'ordre 1 est suffisant pour assurer l'unicité de la solution [Duc76, Buh00].

L'expression analytique d'une telle transformation dans l'espace bidimensionnel est donnée par [Boo89, BZ04] :

$$\begin{cases} T : \mathbb{R}^2 \times \mathbb{R}^{6+2l+2l} & \rightarrow \mathbb{R}^2 \\ (p, \beta) & \mapsto T(p, \beta) = A_{2d}p + t_{2d} + \sum_{k=1}^l w_k \phi(\|p - c_k\|) \end{cases} \quad (2.16)$$

où :

- \*  $\phi$  est une fonction de base,
- \*  $\{c_k\}_{k=1..l}$  sont les centres d'influence autour desquels les fonctions de base sont définies :

$$c_k = [c_k^x \ c_k^y]^T = [\beta_{6+2k-1} \ \beta_{6+2k}]^T$$

- \*  $\{w_k\}_{k=1..l}$  sont les poids affectés aux fonctions de base :

$$w_k = [w_k^x \ w_k^y]^T = [\beta_{6+2l+2k-1} \ \beta_{6+2l+2k}]^T.$$

Dans ce mémoire de thèse, nous utilisons essentiellement l'expression de l'équation (2.16) dans laquelle les fonctions de base sont des TPS.

De telles transformations sont aussi appelées **spline**.

Ces transformations non rigides ont donc deux composantes :

- \* une composante affine qui explique le comportement à l'infini,
- \* et une composante non affine (une fonction multiXXX) qui explique les déformations locales aux centres d'influence dans le cas d'une base de fonctions radiales de type gaussienne ou les déformations loin des centres d'influence dans le cas d'une base de fonctions radiales de type « plaque mince » ou quadratique. Ces dernières déformations peuvent être supérieures au terme affine.

- **transformation par morceaux** : Une transformation non rigide peut être approchée par la déformation d'un maillage (voir chapitre 4). Dans le cas où le maillage est triangulaire, il s'agit d'une transformation affine par morceaux (une transformation affine par triangle). Pour les maillages quadrangulaires, les transformations sont des homographies.

### 2.3.1.3 Exemple d'une image déformée par différents types de transformations

Les figures 2.2 et 2.3 montrent différentes déformations de la première image (figures 2.2-(a) et 2.3-(a)) de la séquence « Foreman ».

Tandis que pour les transformations rigides et affines, les paramètres de ces transformations se trouvent facilement même pour un non-expert (rotation, translation, mise à l'échelle), pour les autres transformations, les paramètres sont plus obscurs et ont donc été estimés en utilisant des méthodes classiques à partir de couples de points  $(\mathbf{m}_j^i, \mathbf{m}_j^f)$ , où  $\mathbf{m}_j^i$  appartient à l'image initiale et  $\mathbf{m}_j^f$  à l'image finale (déformée).

Les figures 2.2-(b), 2.2-(c), 2.2-(d), 2.2-(e), 2.3-(b), 2.3-(c) et 2.3-(d) montrent l'image déformée respectivement par :

- une transformation rigide composée d'une rotation de 30 degrés autour du centre de l'image puis d'une translation de  $[15 \ -5]^T$  pixels.
- une transformation affine composée de la transformation rigide précédente et de la mise à l'échelle de 1.25 selon  $x$  et 0.75 selon  $y$ .
- une homographie. L'homographie normalisée inverse estimée est :

$$H = \begin{bmatrix} 0.6876 & -0.1162 & 49.1854 \\ 0.0318 & 0.6778 & 22.2597 \\ -0.0003 & -0.0005 & 1 \end{bmatrix}$$

- une transformation quadratique. La transformation inverse estimée est paramétrée par :

$$\begin{aligned} \beta(1 : 6) &= [ \ -0.0012 \ 0.0047 \ 0.0014 \ 1.2656 \ -2.2370 \ 171.2766 \ ]^T \\ \beta(7 : 12) &= [ \ -0.0009 \ 0.0019 \ 0.0002 \ 0.4907 \ 0.0808 \ 38.7598 \ ]^T \end{aligned}$$

- une transformation spline de quatre gaussiennes d'écart type  $\sigma = 20$ . Les centres d'influences sont positionnés en :

$$\begin{cases} c_1 = [200 \ 150]^T \\ c_2 = [300 \ 150]^T \\ c_3 = [200 \ 250]^T \\ c_4 = [300 \ 250]^T \end{cases}$$

La transformation inverse estimée est paramétrée par :

$$\begin{aligned} A_{2d} &= \begin{bmatrix} 1.00 & 0.01 \\ 0.29 & 1.24 \end{bmatrix}, t_{2d} = \begin{bmatrix} -4.44 \\ -128.61 \end{bmatrix}, \\ w_1 &= 10^4 \begin{bmatrix} -4.90 \\ 0.61 \end{bmatrix}, w_2 = 10^4 \begin{bmatrix} 5.21 \\ -0.65 \end{bmatrix}, \\ w_3 &= 10^4 \begin{bmatrix} 2.37 \\ -0.29 \end{bmatrix} \text{ et } w_4 = 10^4 \begin{bmatrix} -2.68 \\ 0.33 \end{bmatrix}. \end{aligned}$$

- une transformation spline de quatre splines de type « plaque mince » de mêmes centres d'influence que le réseau de gaussiennes précédent. La transformation inverse estimée est paramétrée par :

$$A_{2d} = \begin{bmatrix} 1.03 & -0.02 \\ 0.28 & 1.25 \end{bmatrix}, t_{2d} = \begin{bmatrix} -52.73 \\ -122.57 \end{bmatrix},$$



(a)



(b)



(c)



(d)



(e)

FIG. 2.2 – Exemples d’une image déformée par différents types de transformation : (a) l’image originale ; (b) transformation rigide (rotation + translation) ; (c) transformation affine (même rotation et translation que la précédente avec ajout d’une mise à l’échelle dans les directions  $x$  et  $y$ ) ; (d) homographie ; (e) transformation quadratique.

$$w_1 = 10^{-3} \begin{bmatrix} -4.67 \\ 0.58 \end{bmatrix}, w_2 = 10^{-3} \begin{bmatrix} 4.98 \\ -0.62 \end{bmatrix},$$

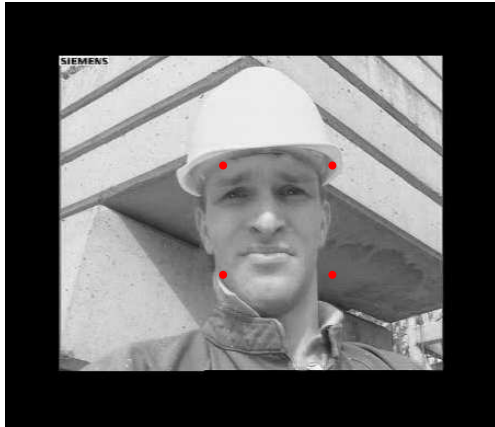
$$w_3 = 10^{-3} \begin{bmatrix} 2.26 \\ -0.28 \end{bmatrix} \text{ et } w_4 = 10^{-3} \begin{bmatrix} -2.56 \\ 0.32 \end{bmatrix}.$$

- une transformation spline de quatre fonctions de type quadratique de paramètre  $a = 100$  et de mêmes centres d'influence que le réseau de gaussiennes précédent. La transformation inverse estimée est paramétrée par :

$$A_{2d} = \begin{bmatrix} 1.05 & -0.04 \\ 0.28 & 1.25 \end{bmatrix}, t_{2d} = \begin{bmatrix} 18.05 \\ -131.42 \end{bmatrix},$$

$$w_1 = \begin{bmatrix} 4.06 \\ -0.50 \end{bmatrix}, w_2 = \begin{bmatrix} -4.32 \\ 0.54 \end{bmatrix},$$

$$w_3 = \begin{bmatrix} -1.96 \\ 0.24 \end{bmatrix} \text{ et } w_4 = \begin{bmatrix} 2.22 \\ -0.27 \end{bmatrix}.$$



(a)



(b)



(c)



(d)

FIG. 2.3 – Exemples d'une image déformée par différents types de transformation : (a) l'image originale et les centres d'influence (en rouge) ; transformation non rigide de fonction de base de type : (b) gaussienne ; (c) « plaque mince » et (d) quadratique.



### 2.3.2 Modèles de variations photométriques

Les variations de niveau de gris d'un pixel peuvent être dues :

- i) au mouvement relatif de la scène et de la caméra : un pixel ne représente plus le même élément de surface tridimensionnel,
- 2i) à la présence d'occultations (voir §2.3.3),
- 3i) à la variation d'éclairage de la scène (déplacement d'une source lumineuse, présence d'ombre, un nuage passant devant le soleil...) et aux propriétés photométriques des surfaces,
- 4i) à la caméra : qualité du capteur CCD (*Charge Coupled Device*), quantification des données.

Tandis que les causes (i) et (2i) donnent plutôt une information relative au mouvement dans une scène, (3i) et (4i) concernent deux images d'un même élément de surface tridimensionnel. Il est à noter que, dans un algorithme de suivi, il est impossible de connaître la cause d'une variation photométrique, donc il est a priori impossible d'en traiter les conséquences de manière indépendante.

Les sections suivantes sont consacrées à un état de l'art de différents modèles de variations photométriques imputables aux causes (3i) et (4i). Nous y notons  $p$  un pixel de l'image courante  $I_t$  et  $p'$  son correspondant dans l'image de référence  $I_{ref}$  ou  $I_{t-1}$ .

#### 2.3.2.1 Invariance photométrique

L'hypothèse d'invariance photométrique indique qu'il y a, dans le temps, conservation des niveaux de gris des images d'un même élément de surface tridimensionnel. Cette hypothèse est valide lorsque la surface est lambertienne, l'éclairage est constant...

Cette hypothèse d'invariance est qualifiée de :

- spatialement :

\* « faible » si elle ne concerne qu'un ensemble  $\mathcal{Z}$  de points cible :

$$\forall p \in \mathcal{Z} \quad (2.17)$$

\* « forte » si elle concerne tous les points :

$$\forall p \in \mathbb{R}^2. \quad (2.18)$$

- temporellement :

\* « faible » si elle concerne deux images consécutives :

$$\forall t > 0, I_t(p) \approx I_{t-1}(p'). \quad (2.19)$$

Elle est à la base de l'équation fondamentale du flot optique [[Hor86], p. 282].

\* « forte » si elle concerne toute la séquence :

$$\forall t > 0, I_t(p) = I_{ref}(p'), \quad (2.20)$$

habituellement,  $I_{ref} = I_0$ . Cette hypothèse est utilisée par [HB98, JD02].

### 2.3.2.2 Modèle affine

Le modèle de variations photométriques le plus simple est le modèle affine utilisé dans [LF95, FM91, Tok98, NL93]. Il est donné par :

$$\forall p, I_t(p) = a(p)I_{ref}(p') + b(p) \quad (2.21)$$

où :

- $a : \mathbb{R}^2 \rightarrow \mathbb{R}$  est un coefficient de contraste,
- $b : \mathbb{R}^2 \rightarrow \mathbb{R}$  est un coefficient de luminosité.

#### A) Le modèle de Li et Forchheimer [LF95] et Fuh et Maragos [FM91]

Les paramètres  $a$  et  $b$  sont constants sur une région  $\mathcal{R}$  [LF95, FM91] :

$$\forall p \in \mathcal{R}, a_{\mathcal{R}}(p) = a_0 \quad (2.22)$$

et

$$\forall p \in \mathcal{R}, b_{\mathcal{R}}(p) = b_0. \quad (2.23)$$

Les fonctions  $a$  et  $b$  ne sont donc pas continues, sauf si les équations (2.22) et (2.23) sont vraies quelque soit  $\mathcal{R}$  :

$$\forall \mathcal{R}, \forall p \in \mathcal{R}, a(p) = a_0 \text{ et } b(p) = b_0. \quad (2.24)$$

#### B) Le modèle de Toklu [[Tok98], éq. (2.8), p. 26]

Candemir Toklu modélise les variations de niveau de gris dans le cadre d'un maillage quadrangulaire. Pour chaque nœud  $n$  du maillage, il associe deux valeurs nodales (voir définition 4.1.1.10) : une valeur de contraste  $a_n$  et de luminosité  $b_n$ . Une interpolation linéaire (cas d'un maillage triangulaire) ou bilinéaire (cas d'un maillage quadrangulaire) est utilisée pour déterminer les valeurs associées à chaque point d'une maille [Tok98], soit :

$$\forall p \in \mathcal{R}, a_{\mathcal{R}}(p) = \sum_{i=1}^3 a_i \lambda_i(p) \quad (2.25)$$

et

$$\forall p \in \mathcal{R}, b_{\mathcal{R}}(p) = \sum_{i=1}^3 b_i \lambda_i(p), \quad (2.26)$$

dans le cas où la région  $\mathcal{R}$  est un triangle du maillage (ici délimité par les nœuds 1, 2 et 3) et où  $\forall i = 1..3, \lambda_i(p)$  sont les coefficients de l'interpolation de type Lagrange de  $p$  dans  $\mathcal{R}$  (voir §4.1.3).

La détermination des contraste  $a_n$  et luminosité  $b_n$  de chaque nœud est effectuée via la résolution d'un problème de minimisation aux moindres carrés linéaires [[Tok98], éq. (3.16)] et est donnée dans [[Tok98], p. 49–51] pour le cas d'un maillage triangulaire. Les fonctions  $a_{\mathcal{R}}$  et  $b_{\mathcal{R}}$  sont continues d'une maille à l'autre.

**C) Le modèle de Nicolas et Labit [NL93]**

Nicolas et Labit [NL93] considèrent qu'il n'y a pas de variation de contraste :

$$\forall p, a(p) = 1, \quad (2.27)$$

et proposent d'approcher les variations de luminosité  $b$  par un modèle polynomial [[NL93], éq. (1), p. 3] :

$$\forall p, I_t(p) = I_{ref}(p') + b(p), \quad (2.28)$$

où [[NL93], éq. (14), p. 11] :

$$\forall p = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{R}, b_{\mathcal{R}}(p) = \sum_{n=0}^k \sum_{i=0}^n b_{ni} (x - x_g)^{n-i} (y - y_g)^i, \quad (2.29)$$

où :

- $\mathcal{R}$  est une région,
- $k$  est le degré de la fonction polynomiale,
- $[x_g \ y_g]^T$  est le centre de gravité de  $\mathcal{R}$ ,
- $\forall n = 1..k, \forall i = 1..n, b_{ni}$  est le  $ni$ -ème coefficient du modèle polynomial. Il est homogène à un niveau de gris.

La fonction  $b$  est donc définie par morceaux (une fonction  $b_{\mathcal{R}}$  par région  $\mathcal{R}$ ). Nicolas et Labit utilisent essentiellement les modèles polynomiaux de degré 0 et 1, c'est-à-dire, respectivement :

$$\forall p \in \mathcal{R}, b_{\mathcal{R}}(p) = b_{00} \quad (2.30)$$

et

$$\forall p \in \mathcal{R}, b_{\mathcal{R}}(p) = b_{00} + b_{10}(x - x_g) + b_{11}(y - y_g). \quad (2.31)$$

Un exemple de détermination des coefficients  $b_{ni}$  dans le cadre d'un modèle polynomial de degré 1 et d'une région rectangulaire est donné par l'ensemble d'équations (29) dans [[NL93], §5.2, p. 16-17]

**2.3.2.3 Modèles fondés sur une base d'apprentissage**

À partir d'une banque de données  $\mathcal{BD}$  (aussi appelée base d'apprentissage) de  $n$  images ( $\mathcal{BD} = \{I^{(1)} = I_{ref}, I^{(2)}, \dots, I^{(n)}\}$ ) d'un même objet prises sous différentes conditions d'éclairage, Hager et Belhumeur [HB98] puis Matthews et Baker [BM01] proposent deux modèles de variations de la forme :

$$A_t = A_0 + B\lambda, \quad (2.32)$$

où :

- $A_0 \in \mathbb{R}^q$  est l'apparence, dite « de référence », de l'objet (autrement dit, la texture qui servira de référence) ;  $q$  est le nombre de pixels de la région d'intérêt  $\Omega_0$  dans l'image de référence  $I_{ref}$ ,
- $A_t \in \mathbb{R}^q$  est l'apparence courante de l'objet,

- $B = [b_1 \dots b_m] \in \mathbb{R}^{q \times m}$  est une matrice dans laquelle  $\mathcal{B} = \{b_i\}_{i=1..m}$  est une base orthonormée des variations d'apparence. Hager et Belhumeur et Matthews et Baker se différencient par la façon de déterminer  $\mathcal{B}$ ,
- $\lambda \in \mathbb{R}^m$  est un vecteur contenant les  $m$  coordonnées des variations exprimées dans la base  $\mathcal{B}$ . Ces coordonnées sont aussi appelées paramètres d'apparence.

**A) Le modèle de Hager et Belhumeur** [[HB98], éq. (41), p. 1031]

Les images de la banque  $\mathcal{BD}$  ont été prises dans les conditions suivantes :

- l'objet est immobile,
- la caméra est fixe,
- les sources lumineuses sont déplacées et leurs intensités varient.

Hager et Belhumeur s'intéressent à toute l'image, donc :

- $q$  est le nombre de pixels de l'image,
- $A_0 = \mathbf{I}_{ref}$ ,
- $A_t = \mathbf{I}_t$ ,
- $B$  est déterminée par l'algorithme 2.3.1.

Soit  $C \in \mathbb{R}^{q \times (n+1)}$ ,  $q > (n+1)$  telle que :

$$C = [C_1 C_2 C_3 \dots C_{n+1}],$$

où :

- $C_1 = \mathbf{I}_{ref}$ ,
- $C_2 = [1 \dots 1]^T$ ,
- $[C_3 \dots C_{n+1}] = [\mathbf{I}^{(2)} \dots \mathbf{I}^{(n)}]$ .

- 1) Calculer une décomposition en valeurs singulières (SVD : *Singular Value Decomposition*) de  $C$  :

$$C = U \Sigma V^T,$$

où :

- $\Sigma \in \mathbb{R}^{(n+1) \times (n+1)}$  est une matrice diagonale formée des valeurs singulières de  $C$  qui sont les racines carrées des valeurs propres de  $C^T C$  (par convention,  $\sigma_1 (= \sigma_{11}) \geq \sigma_2 \geq \dots \geq \sigma_{n+1}$ ),
- $U = [u_1 \dots u_{n+1}] \in \mathbb{R}^{q \times (n+1)}$  est formée de  $n+1$  vecteurs propres orthonormés associés aux  $n+1$  plus grandes valeurs propres de  $C C^T$ ,
- $V = [v_1 \dots v_{n+1}] \in \mathbb{R}^{(n+1) \times (n+1)}$  est formée des  $n+1$  vecteurs propres orthonormés associés aux  $n+1$  valeurs propres de  $C^T C$ .

- 2) Déterminer la base orthonormée  $\mathcal{B}$  telle que :

$$B = [u_1 \dots u_m], m \leq (n+1),$$

c'est-à-dire les  $m$  vecteurs propres associés aux  $m$  plus grandes valeurs singulières de  $C$ .

Algorithme 2.3.1: Construction de la base orthonormée des variations d'intensité lumineuse de Hager et Belhumeur.

**B) Le modèle proposé par Matthews et Baker [BM01]**

Il est utilisé dans le cadre des modèles d'apparence active (AAM : *Active Appearance Model*). Les images de la banque  $\mathcal{BD}$  ont été prises dans les conditions suivantes :

- l'objet est mobile,
- la caméra est mobile,
- les sources lumineuses sont déplacées et leurs intensités varient.

La zone d'intérêt est la partie de l'image dans laquelle l'objet se situe sur l'image de référence. Ils associent un maillage  $\mathcal{M}_0$  à cette partie. Le domaine  $\Omega_0$  représente l'ensemble des pixels de  $\mathcal{M}_0$ . Pour eux :

- $q = |\Omega_0|$ ,
- $A_0$ ,  $A_t$  et  $B$  sont respectivement déterminés par les algorithmes 2.3.2, 2.3.3 et 2.3.4.

– Maillage de référence

**Pour  $i$  de 1 à  $n$  Faire**

$\mathcal{M}_i \leftarrow i\_vers\_m\_main(I^{(i)})$

**Fin Pour**

$\mathcal{M}_0 \leftarrow \frac{1}{n} \sum_{i=1}^n \mathcal{M}_i$

– Apparence de référence

**Pour  $i$  de 1 à  $n$  Faire**

$A^{(i)} \leftarrow i\_m\_m0\_vers\_a(I^{(i)}, \mathcal{M}_i, \mathcal{M}_0)$

**Fin Pour**

$A_0 \leftarrow \frac{1}{n} \sum_{i=1}^n A^{(i)}$

Fonctions utiles :

$i\_vers\_m\_main$  : effectue le maillage de l'objet de façon manuelle

$i\_m\_m0\_vers\_a$  : fournit l'apparence de l'objet aligné géométriquement sur le maillage de référence.

Algorithme 2.3.2: Construction de l'apparence de référence  $A_0$  dans Matthews et Baker.

– Apparence courante

$\mathcal{M}_t \leftarrow i\_vers\_m(I_t)$

$A_t \leftarrow i\_m\_m0\_vers\_a(I_t, \mathcal{M}_t, \mathcal{M}_0)$

Fonction utile :

$i\_vers\_m$  : maillage obtenu manuellement ou en sortie d'un algorithme de suivi voir algorithme 2.3.2

Algorithme 2.3.3: Construction de l'apparence courante  $A_t$  dans Matthews et Baker.

- 1) Centrer les données :  
**Pour**  $i$  **de** 1 **à**  $n$  **Faire**  
 $B^{(i)} \leftarrow A^{(i)} - A_0$   
**Fin Pour**
- 2) Construire  $C$  :  

$$C = [B^{(1)} \dots B^{(n)}] \in \mathbb{R}^{q \times n}.$$
- 3) Calculer la matrice de covariance associée à  $C$  :  

$$S = \frac{1}{n} C^T C,$$
- 4) Calculer l'analyse en composantes principales (ACP) de  $S$  :  

$$S = V D V^T,$$
  
 où :  
  - $D \in \mathbb{R}^{n \times n}$  est une matrice diagonale formée des valeurs propres de  $S$  (par convention,  $\lambda_1 (= \lambda_{11}) \geq \lambda_2 \geq \dots \geq \lambda_n$ ),
  - $V = [v_1 \dots v_n] \in \mathbb{R}^{n \times n}$  est formée des  $n$  vecteurs propres orthonormés associés aux  $n$  valeurs propres.
- 5) Déterminer la base orthonormée  $\mathcal{B}$  telle que :  

$$B = [v_1 \dots v_m], m \leq n,$$
  
 c'est-à-dire les  $m$  vecteurs propres associés aux  $m$  plus grandes valeurs propres.

Algorithme 2.3.4: Construction de la base orthonormée des variations d'intensité lumineuse de Matthews et Baker.

#### 2.3.2.4 Ce que nous en retenons

La texture des objets composant la scène et les conditions dans lesquelles les séquences d'images sont prises peuvent induire des variations photométriques. Lorsqu'une banque d'images de l'objet est disponible, un apprentissage de telles variations peut être effectué, sinon ces variations vont être estimées tout au long du suivi. Pour notre part, dans ce mémoire de thèse, nous nous intéressons exclusivement à l'hypothèse d'invariance photométrique « forte ». Cette hypothèse est évidemment valable pour les séquences que nous traitons.

### 2.3.3 Occultations

#### 2.3.3.1 Définition (Occultation)

On parle d'**occultation** lorsqu'un objet est partiellement ou totalement masqué par un, voire plusieurs objets (déformables ou indéformables, en mouvement ou non).

C'est, par exemple, le cas pour la statue de la séquence d'images Stanislas [[LB00], fig. 1, p. 227] et la vache de la séquence cow [[LB00], fig. 6, p. 230].

### 2.3.3.2 Définition (Auto-occultation)

On parle d'**auto-occultation** lorsque l'objet occultant est une partie de l'objet occulté.

C'est, par exemple, le cas lorsque le bras d'une personne passe devant son visage.

De nombreux chercheurs [AT97a, Dou04, Tok98, GVPG03, GBBS07] se sont intéressés au suivi d'objets occultés. Ce suivi comporte les quatre étapes suivantes :

- détecter le début d'une occultation (les zones masquées, les zones découvertes...) [AT97a],
- suivre correctement l'objet occulté et l'(les) objet(s) occultant(s) pendant l'occultation,
- détecter la fin de l'occultation,
- identifier les objets à la fin de l'occultation.

Gabriel *et al.* [GVPG03] présentent un état de l'art concernant la deuxième étape.

Deux familles de méthodes se distinguent selon qu'elles :

- fusionnent les deux objets lors de l'occultation et les séparent à la fin ;
- considèrent indépendants les deux objets, tout au long du suivi.

---

### Remarque

Dans le cadre de la réalité augmentée, Lepetit et Berger [LB00] proposent une méthode semi-automatique de détection des objets occultants.

---

## 2.4 Suivi

### 2.4.1 Notations

Notons :

- i)  $T$  le modèle de transformation paramétré par  $\beta$ . Il peut être défini par :

$$\left\{ \begin{array}{l} T: \mathbb{R}^2 \times \mathbb{R}^p \rightarrow \mathbb{R}^2 \\ (\mathbf{m}, \beta) \mapsto T(\mathbf{m}, \beta) = \begin{bmatrix} T_1(\mathbf{m}, \beta) \\ T_2(\mathbf{m}, \beta) \end{bmatrix} \end{array} \right. \quad (2.33)$$

Nous supposons que  $\beta_0$ , noté parfois 0, correspond à l'identité donc :

$$\forall \mathbf{m} \in \mathbb{R}^2, \mathbf{m} = T(\mathbf{m}, \beta_0). \quad (2.34)$$

---

**Exemple :** Si  $T$  est une homographie  $H$ , en reprenant la représentation choisie par Matthews et Baker [[BM04], éq. (94), p. 253] :

$$H = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \\ \beta_4 & \beta_5 & \beta_6 \\ \beta_7 & \beta_8 & 1 \end{bmatrix}, \quad (2.35)$$

on a, en notant  $\mathbf{m} = [x \ y]^T$  :

$$T(\mathbf{m}, \beta) = \begin{bmatrix} \frac{(\beta_1+1)x + \beta_2 y + \beta_3}{\beta_7 x + \beta_8 y + 1} \\ \frac{\beta_4 x + (\beta_5+1)y + \beta_6}{\beta_7 x + \beta_8 y + 1} \end{bmatrix} \stackrel{n}{=} \begin{bmatrix} \frac{N_1}{D} \\ \frac{N_2}{D} \end{bmatrix} \quad (2.36)$$

2i)  $T^\beta$  l'application partielle de  $T$  associée à  $\beta$  :

$$\forall \beta \in \mathbb{R}^p, \begin{cases} T^\beta : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ \mathbf{m} \mapsto T^\beta(\mathbf{m}) = T(\mathbf{m}, \beta). \end{cases} \quad (2.37)$$

3i)  $T^{\mathbf{m}}$  l'application partielle de  $T$  associée à  $\mathbf{m}$  :

$$\forall \mathbf{m} \in \mathbb{R}^2, \begin{cases} T^{\mathbf{m}} : \mathbb{R}^p \rightarrow \mathbb{R}^2 \\ \beta \mapsto T^{\mathbf{m}}(\beta) = T(\mathbf{m}, \beta). \end{cases} \quad (2.38)$$

4i)  $\forall t \geq 0$ ,  $\mathbf{I}_t$  le vecteur contenant l'évaluation de l'image  $I_t$  aux points  $\{\mathbf{m}_i\}_{i=1..n}$  transformés par  $T$  :

$$\begin{cases} \mathbf{I}_t : \mathbb{R}^p \rightarrow \mathbb{R}^n \\ \beta \mapsto \mathbf{I}_t(\beta) = \begin{bmatrix} I_t(T(\mathbf{m}_1, \beta)) \\ \vdots \\ I_t(T(\mathbf{m}_n, \beta)) \end{bmatrix} \\ = [I_t(T(\mathbf{m}_i, \beta))]_{i=1..n} \end{cases} \quad (2.39)$$

Remarquons qu'avec les définitions précédentes, on a :

$$\begin{aligned} \mathbf{I}_t(\beta) &= [I_t(T(\mathbf{m}_i, \beta))]_{i=1..n} \\ &= [(I_t \circ T)(\mathbf{m}_i, \beta)]_{i=1..n} \\ &= [(I_t \circ T^\beta)(\mathbf{m}_i)]_{i=1..n} \\ &= [I_t(T^\beta(\mathbf{m}_i))]_{i=1..n} \\ &= [I_t(T^{\mathbf{m}_i}(\beta))]_{i=1..n} \\ &= [(I_t \circ T^{\mathbf{m}_i})(\beta)]_{i=1..n} \end{aligned} \quad (2.40)$$

Nous jonglerons subtilement entre ces différentes notations.

5i)  $R$  le vecteur des résidus :

$$\begin{cases} R : \mathbb{R}^p \rightarrow \mathbb{R}^n \\ \beta \mapsto R(\beta) = \begin{bmatrix} r_1(\mathbf{m}_1, \beta) \\ \vdots \\ r_n(\mathbf{m}_n, \beta) \end{bmatrix} \\ = \mathbf{I}_t(\beta) - \mathbf{I}_0(\beta_0) \end{cases} \quad (2.41)$$

où le  $i$ -ème résidu est :

$$r_i(\mathbf{m}_i, \beta) = (\mathbf{I}_t(\beta))_i - (\mathbf{I}_0(\beta_0))_i \quad (2.42)$$

$$= I_t(T(\mathbf{m}_i, \beta)) - I_0(\mathbf{m}_i) \quad (2.43)$$



- 6i)  $\mathbf{I}'_t(\beta)$  la dérivée de  $\mathbf{I}_t$  évaluée en  $\beta$ . Elle s'identifie (symbole  $\equiv$ ) à la matrice jacobienne  $J_{\mathbf{I}_t}$  de taille  $n \times p$  :

$$\begin{aligned}\mathbf{I}'_t(\beta) &= \left[ (I_t \circ T^{\mathbf{m}_i})'(\beta) \right]_{i=1..n} \\ &= \left[ I'_t(T(\mathbf{m}_i, \beta)) \circ (T^{\mathbf{m}_i})'(\beta) \right]_{i=1..n} \\ &\equiv J_{\mathbf{I}_t}(\beta) \\ &= \left[ (\nabla I_t(T(\mathbf{m}_i, \beta)))^T J_{T^{\mathbf{m}_i}}(\beta) \right]_{i=1..n}\end{aligned}\quad (2.44)$$

où :

- $I'(\mathbf{m})$ , dérivée de  $I$  évaluée en  $\mathbf{m}$ , est identifiée à la matrice jacobienne de taille  $1 \times 2$  :

$$I'(\mathbf{m}) \equiv (\nabla I(\mathbf{m}))^T = \left[ \frac{\partial I}{\partial x}(\mathbf{m}) \quad \frac{\partial I}{\partial y}(\mathbf{m}) \right] \quad (2.45)$$

- $(T^{\mathbf{m}})'(\beta)$ , dérivée de  $T^{\mathbf{m}}$  évaluée en  $\beta$ , est identifiée à la matrice jacobienne de taille  $2 \times p$  :

$$\begin{aligned}(T^{\mathbf{m}})'(\beta) &\equiv J_{T^{\mathbf{m}}}(\beta) \\ &= \begin{bmatrix} \frac{\partial T_1}{\partial \beta_1}(\mathbf{m}, \beta) & \cdots & \frac{\partial T_1}{\partial \beta_p}(\mathbf{m}, \beta) \\ \frac{\partial T_2}{\partial \beta_1}(\mathbf{m}, \beta) & \cdots & \frac{\partial T_2}{\partial \beta_p}(\mathbf{m}, \beta) \end{bmatrix}\end{aligned}\quad (2.46)$$

De même,  $(T^\beta)'(\mathbf{m})$ , dérivée de  $T^\beta$  évaluée en  $\mathbf{m}$ , est identifiée à la matrice jacobienne de taille  $2 \times 2$  :

$$\begin{aligned}(T^\beta)'(\mathbf{m}) &\equiv J_{T^\beta}(\mathbf{m}) \\ &= \begin{bmatrix} \frac{\partial T_1}{\partial x}(\mathbf{m}, \beta) & \frac{\partial T_1}{\partial y}(\mathbf{m}, \beta) \\ \frac{\partial T_2}{\partial x}(\mathbf{m}, \beta) & \frac{\partial T_2}{\partial y}(\mathbf{m}, \beta) \end{bmatrix}\end{aligned}\quad (2.47)$$

- 7i)  $R'(\beta)$  la dérivée de  $R$  évaluée en  $\beta$ . Elle s'identifie à la matrice jacobienne  $J_R$  de taille  $n \times p$ . Or, d'après la définition des résidus  $R$  (équation (2.41)), on a :

$$R'(\beta) \equiv J_R(\beta) = J_{\mathbf{I}_t}(\beta) \quad (2.48)$$

### 2.4.2 Le problème fondamental ( $\mathcal{P}^s$ )

Le problème abordé se limite à retrouver les paramètres  $\beta_t$  de la transformation  $T$  qui permet de superposer une région  $\mathcal{Z}_t$  d'une image (courante de la séquence)  $I_t$  sur son antécédente  $\mathcal{Z}_{ref}$  d'une image  $I_{ref}$  de référence. En général, l'image de référence est l'image précédente  $I_{t-1}$  ou la première image  $I_0$  de la séquence. Pour simplifier les notations, par la suite  $I_{ref} \stackrel{n}{=} I_0$ .

De l'hypothèse d'invariance photométrique forte (§2.3.2.1, équation (2.20)), on en déduit le critère fondamental  $F_t$  :

$$F_t(\beta_t) = \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta_t)) - I_0(T(\mathbf{m}, \beta_0)))^2 \quad (2.49)$$

où  $\mathbf{m}$  est un point de l'ensemble  $\Omega = \{\mathbf{m}_i\}_{i=1..n} \subset \mathcal{Z}_0$ . Il se réécrit :

$$F_t(\beta_t) = \|\mathbf{I}_t(\beta_t) - \mathbf{I}_0(\beta_0)\|_2^2 \quad (2.50)$$

$$= \|R(\beta_t)\|_2^2 \quad (2.51)$$

Le problème  $(\mathcal{P}^{s(uivi)})$  de minimisation de  $F_t$  est un problème aux moindres carrés non linéaires souvent appelé SSD (*Sum of Squared Differences*) :

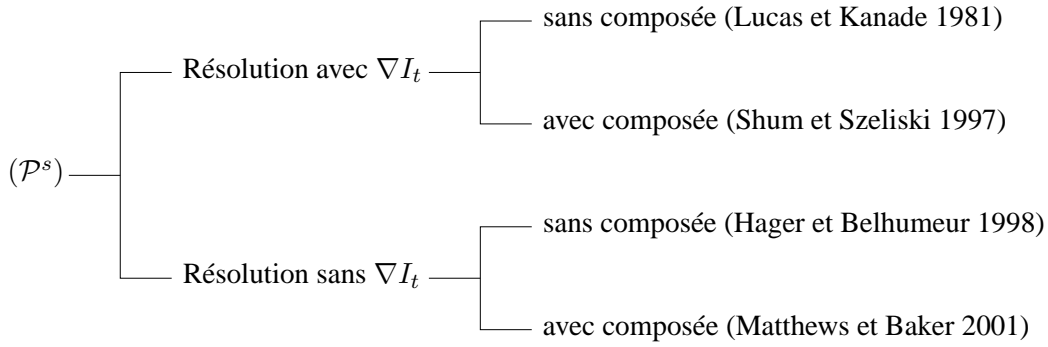
$$(\mathcal{P}^s) \quad \begin{cases} \min & \frac{1}{2} F_t(\beta) \\ \beta \in \mathbb{R}^p \end{cases} . \quad (2.52)$$

Pour l'image  $I_t$  ( $t \geq 1$ ), on doit déterminer  $\beta_t$ , donc  $T^{\beta_t}$ , satisfaisant à :

$$\beta_t = \begin{cases} \operatorname{argmin} & \frac{1}{2} F_t(\beta) \\ \beta \in \mathbb{R}^p \end{cases} . \quad (2.53)$$

### 2.4.3 Quatre approches de résolution de $(\mathcal{P}^s)$

Nous reformulons le cadre unificateur proposé par Matthews et Baker [BM04] à l'occasion du « 20-ième » anniversaire du travail de Lucas et Kanade [LK81], à savoir l'arborescence suivante :



#### 2.4.3.1 Résolution avec $\nabla I_t$ , sans composée

Nommée « *Forwards Additive Approach* » par Matthews et Baker [[BM04], §2, p. 222], il s'agit tout simplement du schéma itératif classique de Gauss-Newton. Rappelons-en les composants à partir des deux interprétations connues [[NW06], §10.3, p. 254–257].

i) Pour une image  $I_t$  ( $t \geq 1$ ) quelconque, le schéma itératif est :

$$\begin{cases} \beta_t^{(0)} \text{ ad-hoc} \\ \beta_t^{(k+1)} \leftarrow \beta_t^{(k)} + s^{gn} \end{cases} , \quad (2.54)$$

où le pas de Gauss-Newton est  $s^{gn}$ .

2i) Première interprétation :

$$s^{gn} = - \left( \widetilde{\nabla^2 F_t(\beta_t^{(k)})} \right)^{-1} \nabla F_t(\beta_t^{(k)}) \quad (2.55)$$

est une approximation, à l'itération  $k$ , du pas de Newton due à l'oubli des termes «  $\nabla^2 r_i$  » dans l'expression de la matrice Hessienne  $\widetilde{\nabla^2 F_t(\beta_t^{(k)})}$ , à savoir :

$$\widetilde{\nabla^2 F_t(\beta_t^{(k)})} = 2 \left( J_R(\beta_t^{(k)}) \right)^T J_R(\beta_t^{(k)}) \quad (2.56)$$

(sachant que  $\nabla F_t(\beta_t^{(k)}) = 2 \left( J_R(\beta_t^{(k)}) \right)^T R(\beta_t^{(k)})$ ).

3i) Seconde interprétation :

En appliquant à  $R$  un développement de Taylor-Young à l'ordre 1 au point  $\beta_t^{(k)}$ ,  $\beta_t^{(k+1)}$  apparaît comme la solution du problème aux moindres carrés linéaires suivant :

$$\begin{cases} \min & \frac{1}{2} \|J_R(\beta_t^{(k)})\beta - (J_R(\beta_t^{(k)})\beta_t^{(k)} - R(\beta_t^{(k)}))\|_2^2, \\ \beta \in \mathbb{R}^p \end{cases}, \quad (2.57)$$

soit :

$$\begin{aligned} \beta_t^{(k+1)} &= \left( (J_R(\beta_t^{(k)}))^T J_R(\beta_t^{(k)}) \right)^{-1} (J_R(\beta_t^{(k)}))^T (J_R(\beta_t^{(k)})\beta_t^{(k)} - R(\beta_t^{(k)})) \\ &= \beta_t^{(k)} - \left( (J_R(\beta_t^{(k)}))^T J_R(\beta_t^{(k)}) \right)^{-1} (J_R(\beta_t^{(k)}))^T R(\beta_t^{(k)}), \end{aligned} \quad (2.58)$$

dès que  $J_R(\beta_t^{(k)})$  est de rang maximal.

Toute la complexité réside dans le calcul de  $J_R(\beta_t^{(k)})$ , qui est donné par l'équation (2.48).

Finalement, on a [[LK81], p. 124] :

$$s^{gn} = (J_{\mathbf{I}_t}(\beta_t^{(k)}))^+ (\mathbf{I}_0(\beta_0) - \mathbf{I}_t(\beta_t^{(k)})) \quad (2.59)$$

### Remarques

- 1) Dans le cas du traitement de l'image  $I_t$  ( $t > 0$ ) d'une séquence  $I_0, I_1, \dots, I_t, \dots$  il va de soi de choisir :

$$\beta_t^{(0)} \text{ (ad-hoc)} = \beta_{t-1}. \quad (2.60)$$

- 2) Il existe des conditions suffisantes assurant la convergence de l'algorithme de Gauss-Newton [[NW06], théorème 10.1, p. 256].
- 3) Matthijs Douze [[Dou04], §4.1.3.2, p. 97] préfère à juste titre résoudre  $(\mathcal{P}^s)$  à l'aide de l'algorithme de Levenberg-Marquardt implémenté dans MINPACK [MGH80].

#### 2.4.3.2 Résolution sans $\nabla I_t$ , sans composée

Nommée « *Inverse Additive Approach* » par Matthews et Baker [[BM04], §3.3, p. 231], il s'agit de la méthode proposée par Hager et Belhumeur [HB98].

Ayant choisi  $\beta_t^{(0)} = \beta_{t-1}$ , le pas de Gauss-Newton  $s^{gn}$  à la première itération :

$$\begin{aligned} s^{gn} &= (J_{\mathbf{I}_t}(\beta_{t-1}))^+ (\mathbf{I}_0(\beta_0) - \mathbf{I}_t(\beta_{t-1})) \\ &= \left[ (\nabla I_t(T(\mathbf{m}_i, \beta_{t-1})))^T J_{T^{\mathbf{m}_i}}(\beta_{t-1}) \right]_{i=1..n}^+ (\mathbf{I}_0(\beta_0) - \mathbf{I}_t(\beta_{t-1})) \end{aligned} \quad (2.61)$$

fait intervenir le gradient de l'image courante  $I_t$  en  $n$  points dépendant de  $\beta_{t-1}$ . Hager et Belhumeur [HB98] font le pari d'utiliser une variante de Gauss-Newton que nous appelons **Gauss-Newton généralisé** par extrapolation du langage de Ciarlet *et al.* [[Cia85], p. 159], en remplaçant :

$$\nabla \mathbf{I}_t(\beta_{t-1}) \text{ par } \nabla \mathbf{I}_{t-1}(\beta_{t-1}).$$

Dans ces conditions, en dérivant les deux membres de l'équation (2.20) de conservation des niveaux de gris, pour  $(t - 1)$ , à savoir :

$$\forall \mathbf{m} \in \mathcal{Z}_0, I_0(\mathbf{m}) = I_{t-1}(T^{\beta_{t-1}}(\mathbf{m})), \quad (2.62)$$

on obtient :

$$\forall \mathbf{m} \in \mathcal{Z}_0, (\nabla I_0(\mathbf{m}))^T = (\nabla I_{t-1}(T(\mathbf{m}, \beta_{t-1})))^T J_{T^{\beta_{t-1}}}(\mathbf{m}) \quad (2.63)$$

soit :

$$\forall \mathbf{m} \in \mathcal{Z}_0, (\nabla I_{t-1}(T(\mathbf{m}, \beta_{t-1})))^T = (\nabla I_0(\mathbf{m}))^T (J_{T^{\beta_{t-1}}}(\mathbf{m}))^{-1} \quad (2.64)$$

et en reportant dans l'expression de  $s^{gn}$ , on a :

$$s^{gn} = \left[ (\nabla I_0(\mathbf{m}_i))^T (J_{T^{\beta_{t-1}}}(\mathbf{m}_i))^{-1} J_{T^{\mathbf{m}_i}}(\beta_{t-1}) \right]_{i=1..n}^+ (\mathbf{I}_0(\beta_0) - \mathbf{I}_t(\beta_{t-1})). \quad (2.65)$$

**Exemple :** Calcul de  $J_{T^\beta}$  dans le cas de l'homographie (équation (2.36)) :

$$J_{T^\beta}(\mathbf{m}) = \frac{1}{D^2} \begin{bmatrix} (\beta_1 + 1)D - \beta_7 N_1 & \beta_2 D - \beta_8 N_1 \\ \beta_4 D - \beta_7 N_2 & (\beta_5 + 1)D - \beta_8 N_2 \end{bmatrix} \quad (2.66)$$

**Exemple :** Calcul de  $J_{T^{\mathbf{m}}}$  dans le cas de l'homographie (équation (2.36)) :

$$J_{T^{\mathbf{m}}}(\beta) = \frac{1}{D} \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -\frac{xN_1}{D} & -\frac{yN_1}{D} \\ 0 & 0 & 0 & x & y & 1 & -\frac{xN_2}{D} & -\frac{yN_2}{D} \end{bmatrix}. \quad (2.67)$$

Notons que Hager et Belhumeur ne proposent à aucun moment d'itérer (contrairement à l'affirmation de Matthews et Baker [[BM04], éq. (45), p. 231]), et qu'interrogés sur la question, ils ont été évasifs en se réfugiant derrière un « c'est bien loin tout ça... ».

Si l'on s'engage sur la piste d'itérer le processus, alors, dès la deuxième itération ( $k \geq 2$ ), la justification peut être différente car :

$$s^{gn} = \left[ (\nabla I_t(T(\mathbf{m}_i, \beta_t^{(k-1)})))^T J_{T^{\mathbf{m}_i}}(\beta_t^{(k-1)}) \right]_{i=1..n}^+ (\mathbf{I}_0(\beta_0) - \mathbf{I}_t(\beta_t^{(k-1)})) \quad (2.68)$$

et il suffit d'appliquer l'équation (2.20) de conservation des niveaux de gris pour  $\beta_t^{(k-1)}$ , à savoir :

$$\forall \mathbf{m} \in \mathcal{Z}_0, I_0(\mathbf{m}) = I_t(T(\mathbf{m}, \beta_t^{(k-1)})). \quad (2.69)$$

Il n'est donc plus nécessaire d'invoquer Newton généralisé et tout repose sur la légitimité de l'équation de conservation des niveaux de gris.

Mentionnons que Ciarlet *et al.* fournissent un théorème établissant les conditions suffisantes de convergence de Newton généralisé [[Cia85], théorème 7.5.1, p. 161]. Transposées à notre cas, ces conditions concernent :

$$\| (J_{\mathbf{I}_{t-1}}(\beta_{t-1}))^T J_{\mathbf{I}_{t-1}}(\beta_{t-1}) \|_{\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)}$$

et

$$\| (J_{\mathbf{I}_{t-1}}(\beta_{t-1}))^T J_{\mathbf{I}_{t-1}}(\beta_{t-1}) - (J_{\mathbf{I}_t}(\beta_{t-1}))^T J_{\mathbf{I}_t}(\beta_{t-1}) \|_{\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)}.$$

### Remarque

En pratique, dans le cas du suivi de motifs au fil des images d'une séquence nous proposons,  $\forall t \geq 1$  :

- i) une itération avec le pas de Gauss-Newton, ce qui oblige à faire, une fois pour chaque image, le calcul explicite de  $\nabla I_t(T(\mathbf{m}_i, \beta_{t-1})), \forall i = 1..n$ .
- 2i) les itérations suivantes selon le schéma proposé par Hager et Belhumeur mais en remplaçant  $\nabla I_0(T(\mathbf{m}_i, \beta_0))$  par  $\nabla I_t(T(\mathbf{m}_i, \beta_{t-1}))$ .

### 2.4.3.3 Résolution avec $\nabla I_t$ , avec composée

Nommée « *Forwards Compositional Approach* » par Matthews et Baker [[BM04], §3.1, p. 225], c'est la solution implicitement proposée par Shum et Szeliski [[SS97], éq. (14), p. 8].

Si l'on fait l'hypothèse que l'ensemble des transformations  $\{T^\beta\}_{\beta \in \mathbb{R}^p}$  muni de la composition ( $\circ$ ) des applications forme un groupe  $\mathcal{T}^\beta = (\{T^\beta\}_{\beta \in \mathbb{R}^p}, \circ)$ , alors l'équation :

$$T_{(k)}^{\beta_t} = T_{(k-1)}^{\beta_t} \circ T^{\Delta\beta} \quad (2.70)$$

a toujours une solution unique en  $T^{\Delta\beta}$  à savoir :

$$T^{\Delta\beta} = \left(T_{(k-1)}^{\beta_t}\right)^{-1} \circ T_{(k)}^{\beta_t} \quad (2.71)$$

[[RDO01a], p. 42]. Remarquons qu'une condition suffisante plus faible est l'existence de  $\left(T_{(k-1)}^{\beta_t}\right)^{-1}$ .

En particulier, pour  $k = 1$  et  $T_{(0)}^{\beta_t} = T^{\beta_{t-1}}$ , on a :

$$T^{\Delta\beta} = \left(T^{\beta_{t-1}}\right)^{-1} \circ T^{\beta_t} \quad (2.72)$$

Dans ces conditions, le critère proposé par Matthews et Baker [[BM04], éq. (12), p. 225] est équivalent au critère initial [[BM04], éq. (3), p. 222].

#### Remarque

Ce n'est qu'au moment de la démonstration de la « pseudo-équivalence », que Matthews et Baker se rendent compte qu'un semi-groupe ne suffit pas [[BM04], p. 228, après l'équation (30)].

Avec nos notations,  $d$  désignant « direct », le critère de l'équation (2.50) devient :

$$F_t^d(\Delta\beta^d) = \|\mathbf{I}_t^d(\Delta\beta^d) - \mathbf{I}_0(\beta_0)\|_2^2 \quad (2.73)$$

où :

$$\begin{cases} \mathbf{I}_t^d : \mathbb{R}^p \rightarrow \mathbb{R}^n \\ \beta \mapsto \mathbf{I}_t^d(\beta) = \left[ \left( I_t \circ T^{\beta_{t-1}} \circ T^\beta \right) (\mathbf{m}_i) \right]_{i=1..n} \\ = \left[ \left( I_t \circ T^{\beta_{t-1}} \circ T^{\mathbf{m}_i} \right) (\beta) \right]_{i=1..n} \end{cases} \quad (2.74)$$

et le problème  $(\mathcal{P}^{s+d})$  à résoudre est :

$$(\mathcal{P}^{s+d}) \quad \left\{ \min_{\beta \in \mathbb{R}^p} \frac{1}{2} F_t^d(\beta) \right. \quad (2.75)$$

Retrouvons les composants du pas de Gauss-Newton à partir de la deuxième interprétation et pour cela, écrivons le développement de Taylor-Young à l'ordre 1 au voisinage de  $\beta_0$  (aussi noté 0) de l'application  $\mathbf{I}_t^d$ .

Intéressons-nous à la composante  $i$  de  $\mathbf{I}_t^d$  :

$$(\mathbf{I}_t^d)_i(0 + \Delta\beta^d) = (\mathbf{I}_t^d)_i(0) + (\mathbf{I}_t^d)'_i(0)\Delta\beta^d \quad (2.76)$$

Il faut donc calculer  $(\mathbf{I}_t^d)'_i(\beta)$ . En convenant de noter  $L(\phi) \stackrel{n}{=} L.\phi \stackrel{n}{=} L\phi$ , pour toute application linéaire  $L$  d'un espace vectoriel  $E$  dans un espace vectoriel  $F$  ( $L \in \mathcal{L}(E, F)$ ) et pour tout vecteur  $\phi$  de  $E$ , on a :

$$\forall \phi \in \mathbb{R}^p, (\mathbf{I}_t^d)'_i(\beta) \phi = I_t'((T^{\beta_{t-1}} \circ T^{\mathbf{m}_i})(\beta)) (T^{\beta_{t-1}} \circ T^{\mathbf{m}_i})'(\beta) \phi \quad (2.77)$$

$$= I_t'(T(T(\mathbf{m}_i, \beta), \beta_{t-1})) ((T^{\beta_{t-1}})' \circ T^{\mathbf{m}_i})(\beta) (T^{\mathbf{m}_i})'(\beta) \phi \quad (2.78)$$

$$= I_t'(T(T(\mathbf{m}_i, \beta), \beta_{t-1})) (T^{\beta_{t-1}})'(T(\mathbf{m}_i, \beta)) (T^{\mathbf{m}_i})'(\beta) \phi \quad (2.79)$$

En identifiant les différentes dérivées avec leurs matrices jacobiniennes respectives, on obtient :

$$(J_{\mathbf{I}_t^d})_i(\beta) = (\nabla I_t(T(T(\mathbf{m}_i, \beta), \beta_{t-1})))^T J_{T^{\beta_{t-1}}}(T(\mathbf{m}_i, \beta)) J_{T^{\mathbf{m}_i}}(\beta). \quad (2.80)$$

En se souvenant que  $T(\mathbf{m}_i, \beta_0) = \mathbf{m}_i, \forall i = 1..n$  (équation (2.34)), soit  $T^{\beta_0} = Id$ , alors :

$$(J_{\mathbf{I}_t^d})_i(\beta_0) = (\nabla I_t(T(\mathbf{m}_i, \beta_{t-1})))^T J_{T^{\beta_{t-1}}}(\mathbf{m}_i) J_{T^{\mathbf{m}_i}}(\beta_0). \quad (2.81)$$

ce qui est l'équivalent de l'équation (17) de Matthews et Baker [BM04].

Matthews et Baker proposent une autre formule fondée sur l'associativité de l'opérateur de composition, à savoir :

$$I_t \circ T^{\beta_{t-1}} \circ T^{\mathbf{m}_i} = (I_t \circ T^{\beta_{t-1}}) \circ T^{\mathbf{m}_i} \quad (2.82)$$

d'où :

$$\forall \phi \in \mathbb{R}^p, (\mathbf{I}_t^d)'_i(\beta) \phi = (I_t \circ T^{\beta_{t-1}})'(T^{\mathbf{m}_i}(\beta)) (T^{\mathbf{m}_i})'(\beta) \phi. \quad (2.83)$$

En identifiant les différentes dérivées avec leurs matrices jacobiniennes respectives, on obtient :

$$(J_{\mathbf{I}_t^d})_i(\beta) = J_{I_t \circ T^{\beta_{t-1}}}(T(\mathbf{m}_i, \beta)) J_{T^{\mathbf{m}_i}}(\beta), \quad (2.84)$$

d'où, pour  $\beta = \beta_0$  et en se souvenant que  $T(\mathbf{m}_i, \beta_0) = \mathbf{m}_i, \forall i = 1..n$  (équation (2.34)), soit  $T^{\beta_0} = Id$  :

$$(J_{\mathbf{I}_t^d})_i(\beta_0) = J_{I_t \circ T^{\beta_{t-1}}}(\mathbf{m}_i) J_{T^{\mathbf{m}_i}}(\beta_0). \quad (2.85)$$

ce qui est l'équivalent de l'équation (19) page 226 de Matthews et Baker [BM04].

Les composants  $A$  et  $b$  pour appliquer le schéma itératif de Gauss-Newton, sur le problème générique  $\begin{cases} \min \\ \beta \in \mathbb{R}^p \end{cases} \frac{1}{2} \|A\beta - b\|_2^2$ , sont :

$$A = [(J_{\mathbf{I}_t^d})_i(\beta_0)]_{i=1..n} \text{ et } b = [I_0(\mathbf{m}_i) - I_t(T(\mathbf{m}_i, \beta_{t-1}))]_{i=1..n} \quad (2.86)$$

---

### Remarque

On n'oubliera pas que  $T^{\beta_t} = T^{\beta_{t-1}} \circ T^{\Delta\beta^d}$  et rien n'empêche d'itérer à condition que :

$$\begin{cases} \beta_t^{(0)} & \leftarrow \beta_{t-1} \\ T_{(k)}^{\beta_t} & \leftarrow T_{(k-1)}^{\beta_t} \circ T^{\Delta\beta^d}, \forall k \geq 1 \end{cases}$$

ce qui revient à dire qu'à chaque itération, on exécute le premier pas du schéma itératif de Gauss-Newton sur un nouveau problème associé à un nouveau critère  $(F_t^d)_{(k)}$  :

$$\begin{cases} (F_t^d)_{(k)}(\Delta\beta^d) = \left\| \left[ (I_t \circ T_{(k-1)}^{\beta_t} \circ T^{\Delta\beta^d})(\mathbf{m}_i) - I_0(\mathbf{m}_i) \right]_{i=1..n} \right\|_2^2 \\ k \geq 1 \end{cases} \quad (2.87)$$

#### 2.4.3.4 Résolution sans $\nabla I_t$ , avec composée

Nommée « *Inverse Compositional Approach* » par Matthews et Baker [[BM04], §3.2, p. 225], c'est leur proposition [[BM04], tableau 6, p. 240].

Il s'agit simplement d'exploiter l'équation fondamentale (2.20) traduisant l'hypothèse d'invariance des niveaux de gris pour transformer le critère  $F_t^d$  (et avec lui le problème  $(\mathcal{P}^{s+d})$ ) en un critère **équivalent**  $F_t^i$  (donc un problème équivalent  $(\mathcal{P}^{s+i})$ ).

Considérons l'image  $I_t$  ( $t \geq 1$ ), alors :

i) sous l'hypothèse où  $\mathcal{T}^\beta = (\{T^\beta\}_{\beta \in \mathbb{R}^p}, \circ)$  est un groupe,

$$\begin{cases} \exists \Delta\beta^d \text{ unique} \\ \exists \Delta\beta^i \text{ unique} \end{cases} \quad \text{satisfaisant à} \quad \begin{cases} T^{\beta_t} = T^{\beta_{t-1}} \circ T^{\Delta\beta^d} \\ T^{\beta_t} = T^{\beta_{t-1}} \circ (T^{\Delta\beta^i})^{-1} \end{cases} \quad ,$$

$$\text{à savoir} \quad \begin{cases} T^{\Delta\beta^d} = (T^{\beta_{t-1}})^{-1} \circ T^{\beta_t} \\ T^{\Delta\beta^i} = (T^{\Delta\beta^d})^{-1} = (T^{\beta_t})^{-1} \circ T^{\beta_{t-1}} \end{cases} \quad .$$

2i) sous l'hypothèse d'invariance des niveaux de gris,  $\forall \mathbf{m} \in \mathbb{R}^2$  :

$$I_0(\mathbf{m}) = I_t(T^{\beta_t}(\mathbf{m})) \quad (2.88)$$

$$\Leftrightarrow I_0(\mathbf{m}) = I_t \left( \left( T^{\beta_{t-1}} \circ (T^{\Delta\beta^i})^{-1} \right) (\mathbf{m}) \right) \quad (2.89)$$

$$\Leftrightarrow I_0 = I_t \circ T^{\beta_{t-1}} \circ (T^{\Delta\beta^i})^{-1} \{ = I_t \circ T^{\beta_{t-1}} \circ T^{\Delta\beta^d} \} \quad (2.90)$$

$$\Leftrightarrow I_0 \circ T^{\Delta\beta^i} = I_t \circ T^{\beta_{t-1}}. \quad (2.91)$$

L'équation (2.90) justifie le critère  $F_t^d$  (équation (2.73)) associé au problème  $(\mathcal{P}^{s+d})$  (*Forwards Compositional*).

L'équation (2.91), équivalente à l'équation (2.90), justifie pour Matthews et Baker le critère  $F_t^i$  [[BM04], éq. (31), p. 228] associé au problème  $(\mathcal{P}^{s+i})$  (*Inverse Compositional*), à savoir :

$$(\mathcal{P}^{s+i}) \quad \left\{ \min_{\beta \in \mathbb{R}^p} \frac{1}{2} F_t^i(\beta) \right\}, \quad (2.92)$$

où :

$$F_t^i(\Delta\beta^i) = \|\mathbf{I}_0^i(\Delta\beta^i) - \mathbf{I}_t^i(\beta_{t-1})\|_2^2 \quad (2.93)$$

avec :

$$\begin{cases} \mathbf{I}_t^i : \mathbb{R}^p \rightarrow \mathbb{R}^n \\ \beta \mapsto \mathbf{I}_t^i(\beta) = \left[ (I_t \circ T^\beta)(\mathbf{m}_j) \right]_{j=1..n} \\ \quad \quad \quad = \left[ (I_t \circ T^{\mathbf{m}_j})(\beta) \right]_{j=1..n} \end{cases} \quad (2.94)$$

**Remarque**

L'équation (2.91) (qui justifie le problème  $(\mathcal{P}^{s+i})$ ) suppose que l'on sait calculer  $I_0$  en tout point  $T^{\Delta\beta^i}(\mathbf{m})$ ,  $\forall \mathbf{m} \in \mathcal{Z}_0$  et aussi  $I_t$  en tout point  $T^{\beta_{t-1}}(\mathbf{m})$ . Dans la pratique, cela conduit à interpoler dans le meilleur des cas et à invalider lorsque le point à évaluer est hors de l'image.

Retrouvons les composants du pas de Gauss-Newton à partir de la deuxième interprétation et pour cela, écrivons le développement de Taylor-Young à l'ordre 1 au voisinage de  $\beta_0$  de l'application  $\mathbf{I}_0^i$ .

Intéressons nous à la composante  $j$  de  $\mathbf{I}_0^i$  :

$$(\mathbf{I}_0^i)_j(0 + \Delta\beta^i) = (\mathbf{I}_0^i)_j(0) + (\mathbf{I}_0^i)'_j(0)\Delta\beta^i \quad (2.95)$$

Il faut donc calculer  $(\mathbf{I}_0^i)'_j(\beta)$ , soit :

$$\forall \phi \in \mathbb{R}^p, (\mathbf{I}_0^i)'_j(\beta) \phi = I_0'(T^{\mathbf{m}_j}(\beta)) (T^{\mathbf{m}_j})'(\beta) \phi. \quad (2.96)$$

En identifiant les différentes dérivées avec leurs matrices jacobienues respectives, on obtient :

$$(J_{\mathbf{I}_0^i})_j(\beta) = (\nabla I_0(T^{\mathbf{m}_j}(\beta)))^T J_{T^{\mathbf{m}_j}}(\beta), \quad (2.97)$$

et, en  $\beta = \beta_0$  :

$$(J_{\mathbf{I}_0^i})_j(\beta_0) = (\nabla I_0(\mathbf{m}_j))^T J_{T^{\mathbf{m}_j}}(\beta_0) \quad (2.98)$$

d'où le critère :

$$F_t^i(\Delta\beta^i) = \left\| \left[ (\nabla I_0(\mathbf{m}_j))^T J_{T^{\mathbf{m}_j}}(\beta_0) \Delta\beta^i + I_0(\mathbf{m}_j) - I_t(T(\mathbf{m}_j, \beta_{t-1})) \right]_{j=1..n} \right\|_2^2 \quad (2.99)$$

Les composants du schéma itératif de Gauss-Newton sont donc :

$$\begin{cases} \ll A \gg = \left[ (\nabla I_0(\mathbf{m}_j))^T J_{T^{\mathbf{m}_j}}(\beta_0) \right]_{j=1..n} & \text{(i)} \\ \ll b \gg = \left[ I_t(T(\mathbf{m}_j, \beta_{t-1})) - I_0(\mathbf{m}_j) \right]_{j=1..n} & \text{(2i)} \\ \ll A^+ \gg = (A^T A)^{-1} A^T & \text{(3i)} \\ \Delta\beta^i = A^+ b & \text{(4i)} \end{cases} \quad (2.100)$$

On constate l'absence du calcul des dérivées sur l'image courante  $I_t$ .

**Remarque**

Comme dans le cas du modèle direct, on peut itérer le processus, c'est-à-dire appliquer à l'itération  $k \geq 1$ , le premier pas d'un algorithme de résolution (Gauss-Newton par exemple) sur un nouveau problème associé à un nouveau critère  $(F_t^i)_{(k)}$  :

$$\begin{cases} \beta_t^{(0)} & \leftarrow \beta_{t-1} \\ T_{(k)}^{\beta_t} & \leftarrow T_{(k-1)}^{\beta_t} \circ (T^{\Delta\beta^i})^{-1} \\ (F_t^i)_{(k)}(\Delta\beta^i) & = \left\| \left[ (I_t \circ T^{\Delta\beta^i})(\mathbf{m}_j) - (I_t \circ T_{(k-1)}^{\beta_t})(\mathbf{m}_j) \right]_{j=1..n} \right\|_2^2 \\ k & \geq 1 \end{cases} \quad (2.101)$$



À propos de l'équivalence des deux problèmes  $(\mathcal{P}^{s+d})$  et  $(\mathcal{P}^{s+i})$  :

- 1) Une condition suffisante pour que  $(\mathcal{P}^{s+d})$  et  $(\mathcal{P}^{s+i})$  aient des solutions respectives  $\Delta\beta^d$  et  $\Delta\beta^i$  conduisant au même  $T^{\beta_t}$  est que les hypothèses d'invariance des niveaux de gris, **spatialement forte** et temporellement forte, s'appliquent. Alors :

$$F_t^d(\Delta\beta^d) = 0 = F_t^i(\Delta\beta^i).$$

- 2) Les hypothèses d'invariance des niveaux de gris spatialement faible et temporellement forte conduisent à  $F_t^d(\Delta\beta^d) = 0$ , mais pas nécessairement à  $F_t^i(\Delta\beta^i) = 0$ , car les niveaux de gris comparés se situent sur  $I_0$  en des points de  $T^{\Delta\beta^i}(\mathcal{Z}_0)$  et sur  $I_t$  en des points  $T^{\beta_{t-1}}(\mathcal{Z}_0)$ .
- 3) À plus forte raison, la condition  $F_t^d(\Delta\beta^d) = 0$  (ou  $F_t^i(\Delta\beta^i) = 0$ ) n'est pas suffisante pour obtenir le même  $T^{\beta_t}$  !
- 4) Il serait vain de faire des spéculations dans un cadre encore moins contraint comme  $F_t^d(\Delta\beta^d)$  petit !

---

#### Remarque

Par le heureux hasard (Taylor-Young à l'ordre 1 au voisinage de  $\Delta\beta^i = 0$ ) des équations du schéma itératif de Gauss-Newton,  $I_0$  est finalement évaluée en les points de  $\mathcal{Z}_0$ , choisis a priori pour avoir des propriétés propices. Il en serait de même pour tout algorithme de la famille « Newton » (Newton généralisé, Quasi-Newton...).

---

#### 2.4.3.5 Bilan

Nous devons retenir que :

- 1) Lucas et Kanade [LK81] résolvent le problème  $(\mathcal{P}^s)$  à l'aide de l'algorithme itératif de Gauss-Newton.
- 2) Hager et Belhumeur [HB98] résolvent le même problème  $(\mathcal{P}^s)$  à l'aide de la première itération d'une version généralisée de Gauss-Newton qui évite ainsi le calcul de  $\nabla I_t$  ( $t \geq 1$ ). Matthews et Baker suggèrent d'itérer le processus sur les mêmes bases.
- 3) Shum et Szeliski [SS97] formulent implicitement un nouveau critère  $(\mathcal{P}^{s+d})$  équivalent à  $(\mathcal{P}^s)$  dès lors que  $\mathcal{T}^\beta = (\{T^\beta\}_{\beta \in \mathbb{R}^p}, \circ)$  est un groupe en décomposant  $T^{\beta_t}$  :

$$T^{\beta_t} = T^{\beta_{t-1}} \circ T^{\Delta\beta^d}. \quad (2.102)$$

Ils résolvent le problème à l'aide de l'algorithme itératif de Gauss-Newton. Notons que tout autre algorithme mieux adapté [[NW06], §10.3 et 10.4, p. 250–266] est utilisable.

- 4) Matthews et Baker [BM04] formulent un nouveau problème  $(\mathcal{P}^{s+i})$  équivalent à  $(\mathcal{P}^{s+d})$  sous l'hypothèse forte d'invariance des niveaux de gris en décomposant  $T^{\beta_t}$  :

$$T^{\beta_t} = T^{\beta_{t-1}} \circ \left(T^{\Delta\beta^i}\right)^{-1}. \quad (2.103)$$

Leur étude comparative de plusieurs algorithmes **itératifs** de résolution [[BM04], §4, p. 240–252] place les algorithmes de Gauss-Newton et de Levenberg-Marquardt (implémentés dans MATLAB) en haut du podium.

## 2.5 Suivi par prédiction linéaire

### 2.5.1 Algorithme de Jurie et Dhome

Dans le cadre du suivi d'objets plans indéformables, Jurie et Dhome [JD02] proposent une méthode de résolution du problème  $(\mathcal{P}^{s+i})$  (équation (2.92)) qui pourrait être qualifiée de cinquième approche.

#### 2.5.1.1 Principe informel

L'algorithme (JD) de suivi de Jurie et Dhome [JD02] se décompose en deux étapes successives :

- une première étape coûteuse mais hors ligne dite **d'apprentissage**, qui repose uniquement sur l'image de référence  $I_0$ . Elle consiste simplement à remplacer avantageusement (plus robuste et plus flexible) :

$$\ll A^+ \gg = [(\mathbf{I}_0^i)'(\beta)]^+ = [(\nabla I_0(\mathbf{m}_j))^T J_{T^{\mathbf{m}_j}}(\beta)]_{j=1..n}^+ \quad (2.104)$$

de l'équation (2.100-(3i)) par l'estimation  $\hat{A}$ .

- une deuxième étape, qui peut être effectuée en ligne, donc en temps réel, dite **de suivi**, consiste à appliquer la relation (2.103) de Matthews et Baker en calculant  $(T^{\Delta\beta^i})^{-1}$  à l'aide du résultat de l'apprentissage.

Cet algorithme s'applique pour divers modèles de mouvement, en particulier pour des transformations homographiques.

#### 2.5.1.2 Étape d'apprentissage

Avant que Matthews et Baker ne publient leurs travaux, Cootes *et al.* [CET01] puis Jurie et Dhome [JD02] ont eu l'idée d'une relation linéaire similaire à l'équation (2.100)-(4i) notée par abus :

$$\Delta\beta = A\Delta\mathbf{I}, \quad (2.105)$$

$$\text{où } \begin{cases} A = \ll A^+ \gg \\ \Delta\mathbf{I} = \ll b \gg \end{cases}.$$

Sur ces bases, Jurie et Dhome choisissent d'estimer  $A$  à partir de la solution du problème aux moindres carrés linéaires suivant :

$$\hat{A} = \begin{cases} \text{argmin} & \sum_{l=1}^{N_a} \|A\Delta\mathbf{I}^l - \Delta\beta^l\|_2^2, \\ A \in \mathbb{R}^{p \times n} \end{cases} \quad (2.106)$$

où :

- $N_a$  est le nombre d'expériences (de perturbations) mises en jeu pour l'apprentissage,
- $\Delta\beta^l \in \mathbb{R}^p$  est la  $l$ -ème perturbation du vecteur de paramètres,
- $\Delta\mathbf{I}^l$  résulte de la perturbation  $\Delta\beta^l$  :

$$\Delta\mathbf{I}^l = \mathbf{I}_0^i(\beta_0 + \Delta\beta^l) - \mathbf{I}_0^i(\beta_0) \in \mathbb{R}^n. \quad (2.107)$$

En notant :

$$\begin{aligned}
B &= [\Delta\beta^1 \dots \Delta\beta^{N_a}] = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} \in \mathbb{R}^{p \times N_a} \\
C &= [\Delta\mathbf{I}^1 \dots \Delta\mathbf{I}^{N_a}] \in \mathbb{R}^{n \times N_a} \\
A &= \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \in \mathbb{R}^{p \times n},
\end{aligned} \tag{2.108}$$

on obtient :  $\hat{A} = BC^+$  [[Bjö96], éq. (9.2.2), p. 343]. Ceci se déduit de :

$$\forall i = 1..p, \hat{a}_i = \left( \underset{a \in \mathbb{R}^n}{\operatorname{argmin}} \quad \| C^T a - b_i^T \|_2^2 \right)^T \tag{2.109}$$

puisque chaque ligne de la matrice  $A$  de la relation linéaire (2.105) conduit à un nouveau second membre «  $b$  » du système linéaire «  $Ax = b$  » alors que «  $A$  » =  $C^T$  est constante.

### 2.5.1.3 Étape de suivi

L'étape de suivi consiste tout simplement à appliquer, d'une image à l'autre, le prédictor linéaire estimé pendant l'étape d'apprentissage, pour en déduire :

i)

$$\begin{aligned}
\Delta\beta &= \hat{A} (\mathbf{I}_t^i(\beta_{t-1}) - \mathbf{I}_0^i(\beta_0)) \\
&= \hat{A} [I_t(T(\mathbf{m}_j, \beta_{t-1})) - I_0(\mathbf{m}_j)]_{j=1..n}
\end{aligned} \tag{2.110}$$

2i)  $T^{\Delta\beta}$

3i)  $(T^{\Delta\beta})^{-1}$

4i)

$$T^{\beta_t} = T^{\beta_{t-1}} \circ (T^{\Delta\beta})^{-1}. \tag{2.111}$$

## 2.5.2 Notre variante (PP)

### 2.5.2.1 Principe informel

Dans le prolongement de l'algorithme (JD), nous proposons une approximation directe des paramètres  $\Delta\beta^-$  de la transformation  $(T^{\Delta\beta})^{-1}$ , élément du groupe  $\mathcal{T}^\beta$ , qui intervient dans l'équation (2.103).

### 2.5.2.2 Étape d'apprentissage (PP\_A)

Nous reprenons l'équation (2.106), où :

- $\hat{A}^-$  remplace  $\hat{A}$  ;
- $T^{-1}$  remplace  $T$  ;
- $(\Delta\mathbf{I}^l)^-$  remplace  $\Delta\mathbf{I}^l$ , où :

$$(\Delta\mathbf{I}^l)^- = [I_0(T^{-1}(\mathbf{m}_j, \beta_0 + \Delta\beta^l) - I_0(\mathbf{m}_j))]_{j=1..n} \tag{2.112}$$

### 2.5.2.3 Étape de suivi (PP\_S)

Nous reprenons l'équation (2.111) en y remplaçant  $(T^{\Delta\beta})^{-1}$  par  $T^{\Delta\beta^-}$ , où :

$$\begin{aligned}\Delta\beta^- &= \hat{A}^- (\mathbf{I}_t^i(\beta_{t-1}) - \mathbf{I}_0^i(\beta_0)) \\ &= \hat{A}^- [I_t(T(\mathbf{m}_j, \beta_{t-1})) - I_0(\mathbf{m}_j)]_{j=1..n}\end{aligned}\quad (2.113)$$

### 2.5.2.4 Ce que nous retenons

- 1) Les calculs des transformations inverses ont uniquement lieu lors de l'apprentissage hors ligne contrairement à (JD).
- 2) Notre proposition se justifie autant que (JD) et joue dans ce contexte un rôle analogue aux « différences finies avant » comparées aux « différences finies arrière » dans le contexte du calcul différentiel approché.
- 3) Les équations (2.110) (JD) et (2.113) (PP) sont toutes les deux fondées sur une évaluation de  $I_t$  en des points identiques, disons  $(\tilde{m}_j)_{j=1..n}$ .

Dans l'apprentissage de (PP), les points de  $I_0$  visités (pour calculer les  $\Delta\mathbf{I}^l$ ) jouent le rôle « d'antécédent » des  $(\tilde{m}_j)_{j=1..n}$ . On ne peut rien dire de tel pour (JD).

## 2.6 Bilan

Dans ce chapitre, nous avons défini les notions permettant d'asseoir notre travail, à savoir :

- qu'est-ce qu'un objet indéformable ou déformable ?
- qu'est-ce qu'une transformation rigide ou non rigide ?
- quelles sont les hypothèses sous-jacentes au suivi (hypothèses géométriques, photométriques) ?

Nous avons posé le problème fondamental ( $\mathcal{P}^s$ ) du suivi et explicité quatre approches de résolution. Ces approches diffèrent selon que le calcul du gradient de l'image  $I_t$  ( $\nabla I_t$ ) est nécessaire ou non et selon qu'il fait intervenir la composition de deux applications ou non, à savoir :

- résolution avec  $\nabla I_t$ , sans composée ;
- résolution sans  $\nabla I_t$ , sans composée ;
- résolution avec  $\nabla I_t$ , avec composée ;
- résolution sans  $\nabla I_t$ , avec composée.

Enfin, nous avons présenté une approche de suivi par prédiction linéaire et notre variante (PP) de ce dernier, qui ne fait pas intervenir le calcul de la transformation inverse lors de l'étape de suivi mais lors de l'étape d'apprentissage.

## Chapitre 3

# Sélection de points d'intérêt pour l'amélioration d'une procédure de suivi d'objets plans indéformables

Les procédures de suivi sont souvent au cœur des algorithmes de vision répondant en temps réel. Il existe une grande variété de **points d'intérêt** qui peuvent être suivis : contours, coins, repères de « sous-images » ou de taches, etc. La sélection de tels points est une étape importante parce qu'elle a une influence directe sur la performance du suivi. Dans la plupart des applications, sans se poser beaucoup de questions, ces points d'intérêt proviennent des détecteurs de coins ou de zones texturées [Mor77, FG87, HS88, ST94, SB95, SB97, Low04, Tri04, TAZ05]. Cependant, tels « bons points d'intérêt » peuvent s'avérer catastrophiques en entrée de tel algorithme de suivi au point de dire a priori : « à chaque algorithme de suivi, ses points d'intérêt ! » [ST94, KH03]. Nous pouvons aussi distinguer les procédures de présélection hors ligne de celles qui évaluent et mettent à jour continuellement l'ensemble de points utilisé pendant le suivi [JFM01, CLL05].

Ce chapitre traite du problème de présélection hors ligne de points d'intérêt dans le cadre de l'amélioration de la procédure de suivi (JD) d'**objets plans indéformables** (définition 2.1.1.1) de Jurie et Dhome [JD02] présentée à la section 2.5.1.

Nos contributions sont :

- 1) un état de l'art détaillé des détecteurs de points d'intérêt axé sur la mise en relief des concepts mathématiques communs et des similitudes et différences entre les détecteurs ;
- 2) l'application au cas de la sélection de points d'intérêt d'approches statistique et algébrique de sélection d'un sous-ensemble ;
- 3) une cascade de filtres permettant la sélection d'un sous-ensemble de points d'intérêt.

Ce chapitre s'articule comme suit :

- un historique des détecteurs de points d'intérêt de 1976 à nos jours ;
- une présentation d'approches statistique et algébrique de sélection d'un sous-ensemble ;
- nos propositions de sélection et leurs applications dans le cadre de l'algorithme (JD) ;
- une étape d'expérimentation ;
- un bilan.

### 3.1 Qu'est-ce qu'un point d'intérêt ?

Dans de nombreuses applications (mise en correspondance d'images stéréoscopiques, détection ou reconnaissance d'objets, suivi...), la comparaison systématique de tous les pixels entre deux images n'est généralement pas envisageable car elle nécessite un temps de calcul très important. Le processus peut être simplifié en réduisant le domaine de recherche aux zones de l'image localement intéressantes. Les points référençant ces zones sont dénommés **points d'intérêt**. De nombreux travaux présentés dans la section 3.2 concernent la détection des points d'intérêt.

Informellement, un point d'intérêt est associé à une discontinuité des niveaux de gris (voire des couleurs), de la texture, de la géométrie, etc. de l'image. Il est souvent assimilé à un coin. Les images de la figure 3.1 donnent une panoplie de types de coins qui se situent à des jonctions de type « L », « V », « T », « Y », « X » et « damier ».

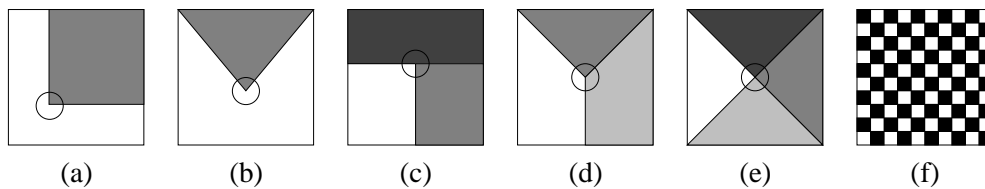


FIG. 3.1 – Différents types de coins : (a) jonction en « L », (b) jonction en « V », (c) jonction en « T », (d) jonction en « Y », (e) jonction en « X » et (f) jonction en « damier ».

### 3.2 Les points d'intérêt et leurs détecteurs au fil du temps

Cette section présente de façon chronologique (voir la frise de la figure 3.2) un catalogue de détecteurs de points d'intérêt publiés depuis 1976.

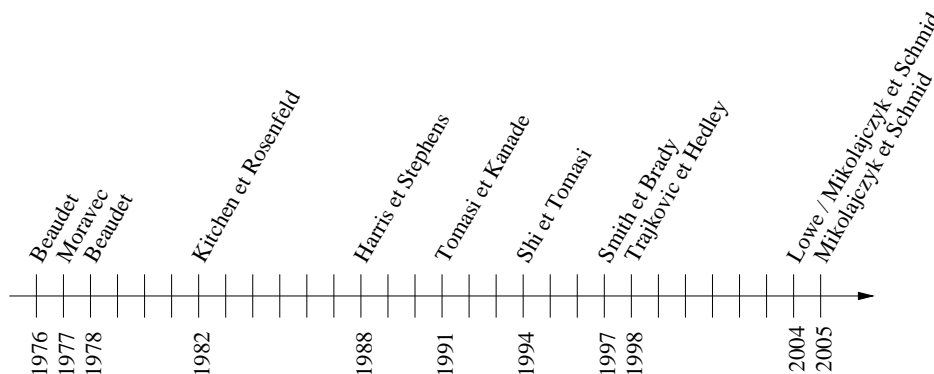


FIG. 3.2 – Frise des détecteurs de points.

L'examen systématique et détaillé de ces détecteurs de points a pour objectif de mettre en lumière leurs fondements « physiques », leurs spécificités propres mais aussi leurs liens.

Une attention toute particulière a été apportée sur l'unification des notations ainsi que sur les modélisations mathématiques sous-jacentes. Nous avons aussi fait en sorte de donner toutes les informations utiles à l'implémentation des algorithmes dédiés. Certaines

sont directement tirées des articles, d'autres sont une proposition d'interprétation des spécifications vagues fournies par les auteurs. Nous avons aussi eu un grand nombre d'échanges électroniques avec différents auteurs. Certaines spécifications en proviennent.

Retenons d'ores et déjà qu'un grand nombre de détecteurs utilise les mêmes outils mathématiques et les mêmes concepts.

Précisons que nous n'avons pas fait d'étude qualitative sur ces détecteurs. Nous verrons dans la section sur l'analyse des résultats que la performance d'un détecteur dépend des images.

Nous commençons par Beaudet [Bea76] en 1976 pour finir avec l'étude collective de Mikolajczyk *et al.* [MTS<sup>+</sup>05] en 2005.

### 3.2.1 Détecteur de Beaudet

#### 3.2.1.1 La théorie

##### A) Le principe informel

L'idée est de rechercher les points-selles (cols) de la fonction image  $I$ . Le pari est de considérer que de tels points, qui ne sont ni minimum ni maximum locaux, révèlent des particularités locales propres aux contours.

##### B) Les équations

Une condition nécessaire d'existence de minimum ou maximum locaux de  $I$  étant :

- i)  $\nabla I(\bar{X}) = 0$
- 2i)  $\nabla^2 I(\bar{X}) \geq 0$  (respectivement  $\leq 0$ ), c'est-à-dire semi-définie positive (respectivement semi-définie négative),

il est donc logique de s'intéresser aux points  $\bar{X} = (\bar{x}, \bar{y})$ , où la matrice hessienne  $\nabla^2 I(\bar{X})$  est indéfinie, ayant donc deux valeurs propres  $\lambda_{min} (< 0)$  et  $\lambda_{max} (> 0)$ . Pour éviter leur calcul, Beaudet [Bea76, Bea78] propose de considérer :

$$\begin{aligned} Det(\nabla^2 I(\bar{X})) &= Det\left(\begin{bmatrix} \frac{\partial^2 I}{\partial x^2}(\bar{X}) & \frac{\partial^2 I}{\partial y \partial x}(\bar{X}) \\ \frac{\partial^2 I}{\partial x \partial y}(\bar{X}) & \frac{\partial^2 I}{\partial y^2}(\bar{X}) \end{bmatrix}\right) \\ &= \frac{\partial^2 I}{\partial x^2}(\bar{X}) \frac{\partial^2 I}{\partial y^2}(\bar{X}) - \left(\frac{\partial^2 I}{\partial x \partial y}(\bar{X})\right)^2 \\ &= \lambda_{min} \lambda_{max} < 0 \end{aligned} \quad (3.1)$$

Il est à noter que, dans son premier article, Beaudet considère «  $-Det$  » [[Bea76], p. 236] ce qui conduit à un résultat positif et une recherche de maxima locaux.

L'effet point-selle (du paraboloïde hyperbolique) est d'autant plus marqué que :

$$\begin{cases} |\lambda_{min}| \approx \lambda_{max} \\ |\lambda_{min}| + \lambda_{max} \text{ est grand.} \end{cases} \quad (3.2)$$

L'exemple suivant, illustré sur la figure 3.3, le montre :

$$\begin{cases} I : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ (x, y) \mapsto I(x, y) = \lambda_{max} x^2 - |\lambda_{min}| y^2, \end{cases} \quad (3.3)$$

où :

- $\lambda_{max} = |\lambda_{min}| = 1$  (figure 3.3-(a)),
- $\lambda_{max} = |\lambda_{min}| = 500$  (figure 3.3-(b)),
- $\lambda_{min} = -1$  et  $\lambda_{max} = 999$  (figure 3.3-(c)).

On recherche donc des minima locaux.

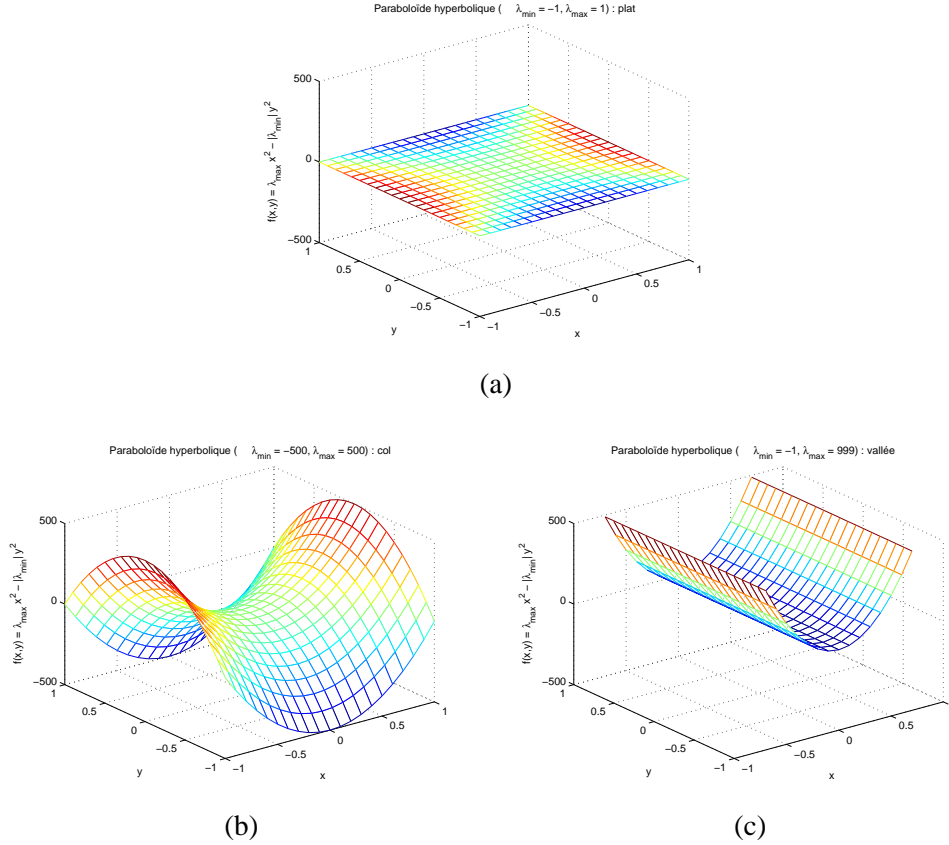


FIG. 3.3 – Paraboloïde hyperbolique : (a) plat ; (b) col ; (c) vallée.

### 3.2.1.2 La pratique

#### A) Calcul approché des dérivées partielles

Au-delà des justifications théoriques, l'efficacité du détecteur de Beudet réside dans le calcul approché des dérivées partielles. Le mérite de Beudet [Bea78] en 1978 est d'avoir rappelé, au bon souvenir de ses contemporains, les polynômes orthogonaux discrets de Tchebychev [Bec73]. Nous faisons une synthèse des tenants et aboutissants de ces polynômes en Annexe A.

Les noyaux de convolution donnés ci-après sont exprimés dans la base canonique  $\{1, x, y, xy, x^2, y^2\}$  [[Bea78], p. 582]. Il est à noter que leurs expressions sont les mêmes dans la base des polynômes orthogonaux discrets de Tchebychev [[Har84], p. 60]. Ces noyaux de convolution permettent d'obtenir les diverses dérivées partielles utiles au calcul du détecteur de Beudet.

- i) Les noyaux de convolution  $3 \times 3$  pour calculer les dérivées partielles jusqu'à l'ordre total de 2 sont les suivants :



$$\frac{\partial I}{\partial x}(\bar{X}) \leftrightarrow \frac{1}{6} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \leftrightarrow \left( \frac{\partial I}{\partial y}(\bar{X}) \right)^T \text{ (Prewitt)} \quad (3.4)$$

$$\frac{\partial^2 I}{\partial x^2}(\bar{X}) \leftrightarrow \frac{1}{3} \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix} \leftrightarrow \left( \frac{\partial^2 I}{\partial y^2}(\bar{X}) \right)^T \quad (3.5)$$

$$\frac{\partial^2 I}{\partial x \partial y}(\bar{X}) \leftrightarrow \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \leftrightarrow \frac{\partial^2 I}{\partial y \partial x}(\bar{X}) \quad (3.6)$$

2i) Les noyaux de convolution  $5 \times 5$  pour calculer les dérivées partielles jusqu'à l'ordre total de 2 sont les suivants :

$$\frac{\partial I}{\partial x}(\bar{X}) \leftrightarrow \frac{1}{50} \begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \leftrightarrow \left( \frac{\partial I}{\partial y}(\bar{X}) \right)^T \quad (3.7)$$

$$\frac{\partial^2 I}{\partial x^2}(\bar{X}) \leftrightarrow \frac{1}{35} \begin{bmatrix} 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \end{bmatrix} \leftrightarrow \left( \frac{\partial^2 I}{\partial y^2}(\bar{X}) \right)^T \quad (3.8)$$

$$\frac{\partial^2 I}{\partial x \partial y}(\bar{X}) \leftrightarrow \frac{1}{100} \begin{bmatrix} 4 & 2 & 0 & -2 & -4 \\ 2 & 1 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 & 2 \\ -4 & -2 & 0 & 2 & 4 \end{bmatrix} \leftrightarrow \frac{\partial^2 I}{\partial y \partial x}(\bar{X}) \quad (3.9)$$

### B) Remarques sur d'autres noyaux de convolution

- **Noyaux de Sobel.** Il n'existe aucun moyen de relier ces polynômes orthogonaux discrets de Tchebychev aux noyaux de convolution de Sobel (voir Danielsson en 1990 [DS90] et réponse de Sobel lui-même, retranscrite en résumé par Danielsson [[DS90], p. 348] et gentiment ressortie de ses archives pour nous (voir Annexe B)).
- **Au sujet du « facet model »** [HW81], [[HS93a], chapitre 8, p. 371]. C'est tout simplement un modèle polynomial d'ordre total égal à 1 qui conduit à des polynômes d'approximation au sens des moindres carrés d'où un problème aux moindres carrés linéaires [[Har84], p. 60], où on retrouve Prewitt.

### C) Minima locaux et seuillage

Il s'agit non seulement de chercher des minima locaux mais en plus de faire précéder cette recherche d'un seuillage double portant sur :

$$\begin{cases} |Det(\nabla^2 I(\bar{X}))| > \tau_{ph}^{(1)} \\ |\lambda_{max} - \lambda_{min}| < \tau_{ph}^{(2)} \end{cases} \quad (3.10)$$

où  $\tau_{ph}^{(1)}$  et  $\tau_{ph}^{(2)}$  ( $ph$  pour paraboloïde hyperbolique) sont les deux seuils. Ceci est une proposition d'implémentation qui traduit, en termes mathématiques, le fait que l'on recherche des zones pour lesquelles la matrice hessienne a des valeurs propres grandes (d'où  $\tau_{ph}^{(1)}$ ) et du même ordre de grandeur (d'où  $\tau_{ph}^{(2)}$ ).

#### D) Concernant la recherche des minima locaux

$\bar{X}$  est minimum local relatif (respectivement strict) de  $f$  si et seulement si :

$$\exists B_{||\cdot||}(\bar{X}, \rho^{(m)}) \forall X (\neq \bar{X}) \in B_{||\cdot||}(\bar{X}, \rho^{(m)}), f(\bar{X}) \leq f(X) \text{ (resp. } < \text{),}$$

où  $B_{||\cdot||}(\bar{X}, \rho^{(m)})$  est la boule ouverte centrée en  $\bar{X}$ , de rayon  $\rho^{(m)}$  ( $(m)$  pour minimum ou maximum) et de norme  $||\cdot||$ .

Lorsque l'on passe en espace discret, cette définition conduit à considérer  $\rho = 1$  et donc  $B_{||\cdot||}(\bar{X}, 1)$ .

Pour Kak et Rosenfeld [[RK76], p. 275 (b)], officiellement à l'origine de la procédure dite de « suppression des points  $\bar{X}$  non maxima locaux », un extremum local  $\bar{X}$  est un extremum sur un voisinage  $B_{||\cdot||}(\bar{X}, \rho^{(m)})$ , où  $\rho^{(m)}$  est prédéfini. Remarquons qu'un point non-maximum local sur  $B_{||\cdot||}(\bar{X}, 3)$  peut l'être sur  $B_{||\cdot||}(\bar{X}, 1)$ .

L'implémentation la plus efficace de cette procédure repose sur une « visite déterministe » des voisins  $X^i$  de  $\bar{X}$  dans un ordre établi selon la distance  $||\bar{X} - X^i||$  croissante.

Ainsi, si  $||\cdot|| = ||\cdot||_2$ , on peut imaginer l'ordre suivant ( $\bar{X}$  étant symbolisé par 0) :

.	.	.	26	.	.	.
.	22	15	10	14	21	.
.	16	6	2	5	13	.
27	11	3	0	1	9	25
.	17	7	4	8	20	.
.	23	18	12	19	24	.
.	.	.	28	.	.	.

et l'algorithme 3.2.1 qui en découle.

```

i ← 1
temp ←  $f(X^{(2\rho^{(m)}+1)^2})$ 
 $f(X^{(2\rho^{(m)}+1)^2}) \leftarrow f(\bar{X}) - 1$  - - on utilise une sentinelle
Tant que ( $f(\bar{X}) \leq f(X^i)$ ) (resp. <)
    i ← i + 1
Fin Tant que
Si ( $i = (2\rho^{(m)} + 1)^2$ ) Alors
     $\bar{X}$  est minimum local
Fin Si
 $f(X^{(2\rho^{(m)}+1)^2}) \leftarrow temp$  - - on rétablit la configuration initiale

```

Algorithme 3.2.1: Algorithme de recherche de minimum local.

### 3.2.2 Détecteur de Moravec

#### 3.2.2.1 La théorie

##### A) Le principe informel

Notons  $\mathcal{Z}(\bar{X}) = B_{\infty}^f(\bar{X}, \rho^{(s)}) \cap \mathbb{Z}^2 \stackrel{n}{=} B_{\infty}^{(d)}(\bar{X}, \rho^{(s)})$ , une zone de l'image  $I$  référencée par le pixel  $\bar{X}$ , où :

- $\rho^{(s)}$  est le rayon de la boule fermée discrète ( $d$ ) correspondant à une valeur de diamètre  $d^{(s)}$  entière et impaire. Par simplification des notations, nous ne prenons pas en compte les diamètres pairs car alors  $\bar{X}$  n'est plus en coordonnées entières et tout se complique.
- ( $s$ ) signifie [auto][dis]similitude.

Si  $\mathcal{Z}(\bar{X})$  est homogène alors l'auto-ressemblance (auto-similitude) locale (c'est-à-dire associée à  $I|_{\mathcal{Z}(\bar{X})}$ ), au sens le plus commun du terme et auto-corrélation pour être plus formel,  $\mathcal{S}^{I|_{\mathcal{Z}(\bar{X})}}$  est maximale de manière isotrope.

Si  $\mathcal{Z}(\bar{X})$  est traversée par une ligne de contours, alors  $\mathcal{S}^{I|_{\mathcal{Z}(\bar{X})}}$  est maximale dans la direction du contour.

Moravec [Mor77] propose de retenir comme zones d'intérêt, et en conséquence comme points d'intérêt, les  $\mathcal{Z}(\bar{X})$  associées à des valeurs :

$$\left\{ \begin{array}{l} \max \\ \phi \in \mathcal{Z}(\bar{X}) \end{array} \mathcal{S}^{I|_{\mathcal{Z}(\bar{X})}}(\phi) \right.$$

faibles relativement aux  $\bar{X}^j$  voisins.

##### B) Les équations

Introduisons la notation suivante :  $[0, N-1]_{\mathbb{N}} \stackrel{n}{=} [0, N-1] \cap \mathbb{N}$ .

Soit  $f : [0, N-1]_{\mathbb{N}} \rightarrow \mathbb{R}$ , appelons  $f^p : \mathbb{Z} \rightarrow \mathbb{R}$  l'extension de période  $N$  de  $f$  satisfaisant à :

$$\left\{ \begin{array}{ll} f^p(i+N) = f^p(i), & \forall i \in \mathbb{Z} \\ f^p(i) = f(i), & \forall i \in [0, N-1]_{\mathbb{N}}. \end{array} \right. \quad (3.11)$$

Alors la fonction d'autocorrélation  $\mathcal{S}^f$  associée à  $f$  ([Bra86], p. 368] et [[AR75], p. 137]) est définie par :

$$\left\{ \begin{array}{ll} \mathcal{S}^f : [0, N-1]_{\mathbb{N}} & \rightarrow \mathbb{R} \\ i & \mapsto \mathcal{S}^f(i) = \sum_{j=0}^{N-1} f^p(i) f^p(i+j). \end{array} \right. \quad (3.12)$$

Considérons la fonction de dissemblance (auto-dissemblance ici)  $\mathcal{D}^f$  associée à  $f$  définie par :

$$\left\{ \begin{array}{ll} \mathcal{D}^f : [0, N-1]_{\mathbb{N}} & \rightarrow \mathbb{R} \\ i & \mapsto \mathcal{D}^f(i) = \sum_{j=0}^{N-1} (f^p(i) - f^p(i+j))^2. \end{array} \right. \quad (3.13)$$

Alors les points qui maximisent  $\mathcal{S}^f$  coïncident avec ceux qui minimisent  $\mathcal{D}^f$  et réciproquement puisque :

$$\mathcal{D}^f(i) = 2 \left( \sum_{j=0}^{N-1} (f(j))^2 - \mathcal{S}^f(i) \right). \quad (3.14)$$

Moravec propose d'utiliser l'extension au cas bidimensionnel de la fonction d'auto-dissemblance  $\mathcal{D}^{I_{\mathcal{Z}(X)}}$  en restreignant la recherche du minimum aux vecteurs :

$$\phi \in \Phi^{\mathcal{M}} = \left\{ \begin{bmatrix} 1 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 1 \end{bmatrix}^T, \begin{bmatrix} 0 & 1 \end{bmatrix}^T, \begin{bmatrix} -1 & 1 \end{bmatrix}^T \right\} \quad (3.15)$$

parmi tous les vecteurs  $\begin{bmatrix} i & j \end{bmatrix}^T$  issus de  $\mathbb{Z} \times \mathbb{Z}$  :

$$\begin{bmatrix} i & j \end{bmatrix}^T \in \left[ -\frac{d^{(s)}+1}{2}, \frac{d^{(s)}+1}{2} \right]_{\mathbb{N}} \times \left[ -\frac{d^{(s)}+1}{2}, \frac{d^{(s)}+1}{2} \right]_{\mathbb{N}}, \text{ si } d^{(s)} \text{ est impair} \quad (3.16)$$

et

$$\begin{bmatrix} i & j \end{bmatrix}^T \in \left[ -\frac{d^{(s)}}{2}, \frac{d^{(s)}}{2} \right]_{\mathbb{N}} \times \left[ -\frac{d^{(s)}}{2}, \frac{d^{(s)}}{2} \right]_{\mathbb{N}} \text{ si } d^{(s)} \text{ est pair,} \quad (3.17)$$

pour définir une application  $\mathcal{M}$  :

$$\begin{cases} \mathcal{M} : \mathbb{N}_n \times \mathbb{N}_m \rightarrow \mathbb{R} \\ X = (x, y) \mapsto \mathcal{M}(X) = \begin{cases} \min_{\phi \in \Phi^{\mathcal{M}}} \mathcal{D}^{I_{\mathcal{Z}(X)}}(\phi) \end{cases} \end{cases} \quad (3.18)$$

où  $\mathbb{N}_k \stackrel{n}{=} [1, k]_{\mathbb{N}}$ .

Les éléments utiles à l'évaluation de cette application  $\mathcal{M}$ , dans le cas simple d'une image de synthèse, sont illustrés sur la figure 3.4.

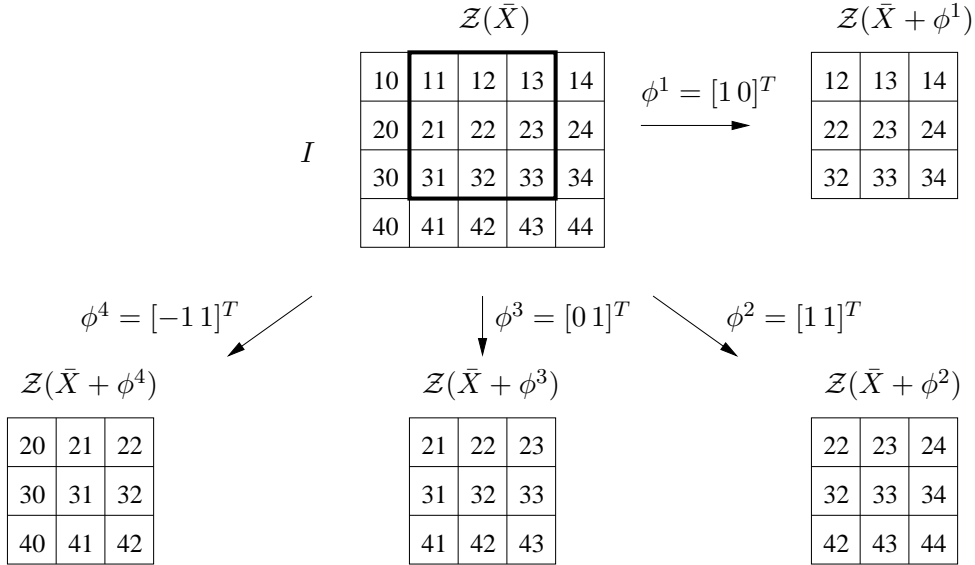


FIG. 3.4 – Exemple d'une image  $I$  avec une zone  $\mathcal{Z}(\bar{X}) = B_{\infty}^{(d)}([3 \ 2]^T, 1)$  et ses décalages  $\mathcal{Z}(\bar{X} + \phi)$ , où  $\phi \in \Phi^{\mathcal{M}}$ .

$\bar{X}$  est « élu » point d'intérêt si et seulement si :

$$\begin{cases} \bar{X} = \begin{cases} \operatorname{argmax} \\ X \in B_{\|\cdot\|}(\bar{X}, \rho^{(m)}) \end{cases} \mathcal{M}(X) \\ \text{et} \\ \mathcal{M}(\bar{X}) > \tau^{\mathcal{D}} \end{cases} \quad (3.19)$$

Le seuil  $\tau^{\mathcal{D}}$  permet de supprimer des points d'intérêt non significatifs.

### 3.2.2.2 La pratique

- A) Moravec a fait des implémentations pour  $d^{(s)} = 3, \dots, 7$
- B) Il retient les maxima locaux à la boule discrète fermée  $B_\infty^{(d)}(\bar{X}, \rho^{(m)} = 1)$ .
- C) Des améliorations ont été proposées pour prendre en considération les vecteurs :

$$\left\{ \begin{bmatrix} 1 & -1 \end{bmatrix}^T, \begin{bmatrix} -1 & 0 \end{bmatrix}^T, \begin{bmatrix} -1 & -1 \end{bmatrix}^T, \begin{bmatrix} 0 & -1 \end{bmatrix}^T \right\}.$$

- D) Naturellement, la phase d'extension périodique de  $I|_{\mathcal{Z}(X)}$  n'a pas lieu puisque  $I(X + \phi)$  est toujours définie tant qu'on ne « déborde » pas de l'image.

### 3.2.2.3 Remarques

#### A) Concernant la fonction de dissemblance $\mathcal{D}^{I|_{\mathcal{Z}(\bar{X})}}$

L'expression de l'équation (3.14) montre que la recherche des points d'intérêt  $\bar{X}$ , comme maxima locaux, est faussée par le niveau de gris moyen de la région d'intérêt.

Un remède à cette anomalie consisterait à recentrer les niveaux de gris autour de cette moyenne mais alors, d'un point de vue complexité, l'intérêt serait moindre.

#### B) Concernant la précision de la localisation des points d'intérêt

Elle diminue en fonction de  $\rho^{(s)}$ , comme le montrent les figures 3.5-(b), (c) et (d), où nous avons simulé l'application de l'opérateur de Moravec à l'image de synthèse de la figure 3.5-(a) avec une recherche de minima locaux relatifs.

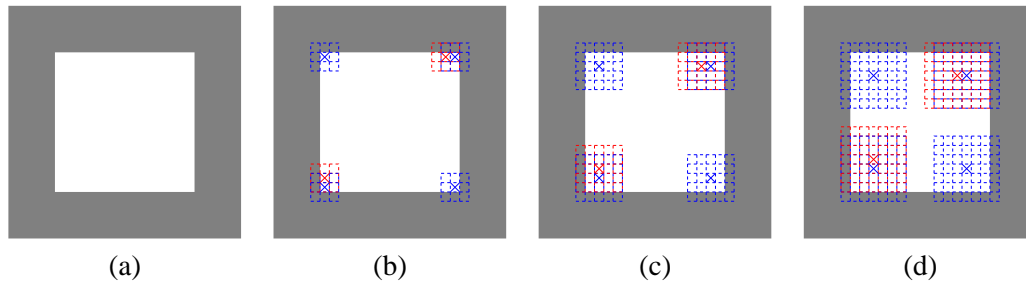


FIG. 3.5 – Position du coin  $\bar{X}$  détecté, sur l'image de synthèse (a), (symbolisé par la croix) par le détecteur de Moravec en fonction de  $\rho^{(s)}$  : (b)  $\rho^{(s)} = 1$ , (c)  $\rho^{(s)} = 2$  et (d)  $\rho^{(s)} = 3$ . Plus la dimension est grande, plus le coin détecté est loin du coin réel.

#### C) Concernant l'isotropie de l'opérateur de Moravec

L'opérateur n'est pas isotrope. La figure 3.5 en est une illustration. Selon les coins, l'opérateur de Moravec peut avoir des réponses différentes. Ici, deux réponses pour les coins haut-droit et bas-gauche.

### 3.2.3 Détecteur de Kitchen et Rosenfeld

#### 3.2.3.1 La théorie

##### A) Le principe informel

Le pari de Kitchen et Rosenfeld [KR82] tient dans la condition nécessaire suivante : « Si un point  $\bar{X}$  appartenant à une ligne de contours est un coin, alors localement, la surface équipotentielle ( $\{X | I(X) = I(\bar{X})\}$ ) passant par  $\bar{X}$  est une courbe plane de très forte courbure algébrique (en valeur absolue) en  $\bar{X}$ , relativement aux points voisins de cette courbe ».

##### B) Les équations

Soit :

$$\begin{cases} P : B_{||\cdot||}(\bar{X}, \rho) \rightarrow \mathbb{R} \\ X = [x \ y]^T \mapsto P(X) \end{cases} \quad (3.20)$$

une fonction polynomiale à deux indéterminées interpolant (ou approchant au sens des moindres carrés) l'image  $I$  sur un voisinage de  $\bar{X}$ . Alors, en supposant satisfaites les hypothèses du théorème des fonctions implicites [[RDO98a], p. 349], l'équation :

$$Q(X) = P(X) - P(\bar{X}) = 0 \quad (3.21)$$

permet de définir une application  $\psi$  :

$$\begin{cases} \psi : \Omega_{\bar{x}} \subset \mathbb{R} \rightarrow \Omega_{\bar{y}} \subset \mathbb{R} \\ x \mapsto \psi(x) = y \end{cases} \quad (3.22)$$

de classe  $\mathcal{C}^\infty$ , où  $\Omega_{\bar{x}}$  (respectivement  $\Omega_{\bar{y}}$ ) est un intervalle ouvert contenant  $\bar{x}$  (respectivement  $\bar{y}$ ), et sa dérivée est donnée par :

$$\psi'(x) = - \frac{\frac{\partial Q}{\partial x}(x, \psi(x))}{\frac{\partial Q}{\partial y}(x, \psi(x))} \quad (3.23)$$

---

##### Remarque

L'application  $Q$  étant polynomiale, elle est de classe  $\mathcal{C}^\infty$  et alors la condition suffisante pour appliquer le théorème des fonctions implicites est :

$$\frac{\partial Q}{\partial y}(\bar{X}) \neq 0 \quad (3.24)$$

---

L'expression de la courbure algébrique d'une courbe plane (voir Annexe C), en  $\bar{X} = (\bar{x}, \bar{y})$ , est alors [[RDO98b], p. 106] :

$$\kappa_c(\bar{X}) = \frac{\psi''(\bar{x})}{(1 + \psi'(\bar{x})^2)^{3/2}} \quad (3.25)$$

Notons :

$$P_x = \frac{\partial P}{\partial x}(\bar{X}) \quad P_y = \frac{\partial P}{\partial y}(\bar{X})$$

$$P_{xx} = \frac{\partial^2 P}{\partial x^2}(\bar{X}) \quad P_{yy} = \frac{\partial^2 P}{\partial y^2}(\bar{X})$$

$$P_{xy} = P_{yx} = \frac{\partial^2 P}{\partial x \partial y}(\bar{X}) = \frac{\partial^2 P}{\partial y \partial x}(\bar{X})$$

Alors  $\psi'(\bar{x}) = -\frac{P_x}{P_y}$  et il reste à calculer  $\psi''(\bar{x})$ .

Soit :

$$\left\{ \begin{array}{lcl} \psi' : \mathbb{R} & \xrightarrow{\psi_1} & \mathbb{R}^2 & \xrightarrow{\psi_2} & \mathbb{R} \\ x & \longmapsto & (x, \psi(x)) = (u, v) & \longmapsto & -\frac{\frac{\partial P}{\partial x}(u, v)}{\frac{\partial P}{\partial y}(u, v)}, \end{array} \right. \quad (3.26)$$

alors, on a :

$$\forall \phi \in \mathbb{R}, \psi'_1(x)\phi = (\phi, \psi'(x)\phi) \quad (3.27)$$

et :

$$\begin{aligned} \forall [\phi_1 \ \phi_2]^T \in \mathbb{R}^2, \psi'_2(u, v)(\phi_1, \phi_2) &= \frac{\partial \psi_2}{\partial x}(u, v)\phi_1 + \frac{\partial \psi_2}{\partial y}(u, v)\phi_2 \\ &= \frac{-\frac{\partial^2 P}{\partial x^2}(u, v)\frac{\partial P}{\partial y}(u, v) + \frac{\partial^2 P}{\partial x \partial y}(u, v)\frac{\partial P}{\partial x}(u, v)}{\left(\frac{\partial P}{\partial y}(u, v)\right)^2}\phi_1 \\ &\quad + \frac{-\frac{\partial^2 P}{\partial y \partial x}(u, v)\frac{\partial P}{\partial y}(u, v) + \frac{\partial^2 P}{\partial y^2}(u, v)\frac{\partial P}{\partial x}(u, v)}{\left(\frac{\partial P}{\partial y}(u, v)\right)^2}\phi_2 \end{aligned} \quad (3.28)$$

Donc, pour l'application composée :

$$\begin{aligned} \forall \phi \in \mathbb{R}, \psi''(x)\phi &= \psi'_2(\psi_1(x))(\psi'_1(x)\phi) \\ &= \psi'_2(x, \psi(x))(\phi, \psi'(x)\phi) \\ &= \psi'_2(x, y)(\phi, \psi'(x)\phi) \\ &= \frac{-\frac{\partial^2 P}{\partial x^2}(x, y)\frac{\partial P}{\partial y}(x, y) + \frac{\partial^2 P}{\partial x \partial y}(x, y)\frac{\partial P}{\partial x}(x, y)}{\left(\frac{\partial P}{\partial y}(x, y)\right)^2}\phi \\ &\quad + \frac{-\frac{\partial^2 P}{\partial x \partial y}(x, y)\frac{\partial P}{\partial y}(x, y) + \frac{\partial^2 P}{\partial y^2}(x, y)\frac{\partial P}{\partial x}(x, y)}{\left(\frac{\partial P}{\partial y}(x, y)\right)^2}\psi'(x)\phi. \end{aligned} \quad (3.29)$$

D'où en utilisant les notations abrégées et en identifiant  $\psi''(\bar{x})$  à la matrice Hessienne  $\nabla^2 \psi(\bar{x})$  :

$$\begin{aligned} \psi''(\bar{x}) &\equiv \nabla^2 \psi(\bar{x}) \\ &= \frac{-P_{xx}P_y + P_{xy}P_x}{P_y^2} - \frac{-P_{xy}P_y + P_{yy}P_x}{P_y^2} \frac{P_x}{P_y} \\ &= -\frac{P_{xx}P_y^2 + P_{yy}P_x^2 - 2P_{xy}P_xP_y}{P_y^3} \end{aligned} \quad (3.30)$$

d'où l'expression de la courbure algébrique :

$$\kappa_c(\bar{X}) = -\frac{P_{xx}P_y^2 + P_{yy}P_x^2 - 2P_{xy}P_xP_y}{(P_x^2 + P_y^2)^{3/2}} \quad (3.31)$$

Kitchen et Rosenfeld [[KR82], p. 98] proposent une normalisation à partir de la norme du vecteur gradient  $\nabla P(\bar{X}) = [P_x \ P_y]^T$ , d'où le critère :

$$\begin{aligned} \mathcal{KR}(\bar{X}) &= (P_x^2 + P_y^2)^{1/2} |\kappa_c(\bar{X})| \\ &= \frac{|P_{xx}P_y^2 + P_{yy}P_x^2 - 2P_{xy}P_xP_y|}{P_x^2 + P_y^2} \end{aligned} \quad (3.32)$$

### 3.2.3.2 La pratique

- A) Pour approcher les dérivées partielles, Kitchen et Rosenfeld utilisent le modèle polynomial à deux indéterminées et de degré total égal à 2 avec des noyaux de convolution  $3 \times 3$ ,  $5 \times 5$  et  $7 \times 7$ . Dans ce dernier cas, selon Beaudet [[Bea78], p. 582], les noyaux de convolution exprimés dans la base canonique sont :

$$\frac{\partial I}{\partial x}(\bar{X}) \leftrightarrow \frac{1}{196} \begin{bmatrix} -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{bmatrix} \leftrightarrow \frac{\partial I}{\partial y}(\bar{X})^T \quad (3.33)$$

$$\frac{\partial^2 I}{\partial x^2}(\bar{X}) \leftrightarrow \frac{1}{294} \begin{bmatrix} 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \end{bmatrix} \leftrightarrow \frac{\partial^2 I}{\partial y^2}(\bar{X})^T \quad (3.34)$$

$$\frac{\partial^2 I}{\partial x \partial y}(\bar{X}) \leftrightarrow \frac{1}{784} \begin{bmatrix} 9 & 6 & 3 & 0 & -3 & -6 & -9 \\ 6 & 4 & 2 & 0 & -2 & -4 & -6 \\ 3 & 2 & 1 & 0 & -1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -6 & -4 & -2 & 0 & 2 & 4 & 6 \\ -9 & -6 & -3 & 0 & 3 & 6 & 9 \end{bmatrix} \leftrightarrow \frac{\partial^2 I}{\partial y \partial x}(\bar{X}) \quad (3.35)$$

En passant au degré total égal à 3, rien ne change pour les noyaux  $3 \times 3$ , par contre pour les noyaux  $5 \times 5$  et  $7 \times 7$  exprimés dans la base canonique :

$$\{1, x, y, xy, x^2, y^2, xy^2, yx^2, x^3, y^3\},$$

$\frac{\partial I}{\partial x}(\bar{X})$  et donc  $\frac{\partial I}{\partial y}(\bar{X})$  sont modifiés.

- B) Ne serait-il pas préférable de conserver la valeur absolue de la courbure après une étape de seuillage sur la norme de gradient ?

### 3.2.3.3 Remarques

#### A) Concernant la variante de Wang et Brady [WB95]

Ils remplacent la courbure  $\mathcal{KR}$  (équation (3.32)) par la courbure moyenne  $\kappa_m$  (voir Annexe C) de la surface définie par l'équation  $P(\bar{x}, \bar{y}) - \bar{z} = 0$ , c'est-à-dire [[WB95], éq. (5), p. 697] :

$$\kappa_m(\bar{X}) = \frac{(1 + P_x^2)P_{yy} + (1 + P_y^2)P_{xx} - 2P_x P_y P_{xy}}{(1 + P_x^2 + P_y^2)^{3/2}} \quad (3.36)$$



[[RDO98b], éq. (8), p. 190].

Ainsi, lorsque  $\|\nabla P(\bar{X})\|_2 \gg 1$ , on retrouve  $\kappa_m(\bar{X}) \approx \kappa_c(\bar{X})$  et la formule de Kitchen et Rosenfeld.

En pratique, l'implémentation de Wang et Brady se distingue de celle de Kitchen et Rosenfeld à trois niveaux :

- i) L'image brute  $I$  est lissée par convolution avec un noyau gaussien 2D centré  $G_\sigma$ , soit :

$$\tilde{I} = I \otimes G_\sigma. \quad (3.37)$$

- 2i) Un seuillage est effectué à partir des valeurs de  $\tilde{\kappa}_m$ . Seuls les points  $\bar{X}$  satisfaisant à :

$$(\tilde{\kappa}_m(\bar{X}))^2 > \tau_{\tilde{\kappa}_m} \quad (3.38)$$

sont conservés.

- 3i) Un second seuillage est effectué sur la norme du gradient de l'image lissée, soit :

$$\|\nabla \tilde{I}(\bar{X})\|_2 > \tau_g. \quad (3.39)$$

Au final, Wang et Brady retiennent les points  $\bar{X}$  satisfaisant à [[WB95], éq. (9), p. 698] :

$$\left\{ \begin{array}{l} \|\nabla \tilde{I}(\bar{X})\|_2 > \tau_g \\ (\tilde{\kappa}_m(\bar{X}))^2 > \tau_{\tilde{\kappa}_m} \\ \frac{\tilde{P}_x^2 \tilde{P}_{xx} + 2\tilde{P}_x \tilde{P}_y \tilde{P}_{xy} + \tilde{P}_y^2 \tilde{P}_{yy}}{\tilde{P}_x^2 + \tilde{P}_y^2} \approx 0 \\ \bar{X} = \begin{cases} \text{argmax} & ((\tilde{\kappa}_m(\bar{X}))^2 - \tau_{\tilde{\kappa}_m}) \\ X \in B_{\|\cdot\|}^{(d)}(\bar{X}, \{1, 2, \dots\}) \end{cases} \end{array} \right. \quad (3.40)$$

## B) Concernant la variante de Zheng *et al.* [ZWT99]

Prétextant que certaines approximations de  $P_{xy}$  sont nulles (ce qui est faux avec les noyaux de Beaudet [[Bea78], p. 582–583]), Zheng *et al.* proposent [[ZWT99], éq. (9), p. 155] :

$$ZWT(\bar{X}) = P_x^2 P_{yy}^2 + P_y^2 P_{xx}^2 - K_{ZWT}(\bar{X}) \|\nabla P(\bar{X})\|_2^4, \quad (3.41)$$

avec :

$$K_{ZWT}(\bar{X}) = \left( G_1 \otimes \frac{P_x^2 P_{yy}^2 + P_y^2 P_{xx}^2}{\|\nabla P(\bar{X})\|_2^4} \right) (\bar{X}), \quad (3.42)$$

où  $G_1$  est un noyau gaussien d'écart type 1.

### 3.2.4 Détecteur de Harris

#### 3.2.4.1 La théorie

##### A) Le principe informel

La volonté de Harris et Stephens [[HS88], p. 149] est d'améliorer l'opérateur de Moravec [Mor77] en palliant trois de ses insuffisances, à savoir :

- i) l'anisotropie ;
- 2i) la sensibilité au bruit ;

3i) la sensibilité aux points de contours qui ne sont pas des coins.

Les remèdes proposés consistent à :

- 1i) étendre  $\Phi^{\mathcal{M}}$  non seulement à  $B_{\infty}^{(d)}(O, 1)$  mais carrément à  $B_2^{(f)}(O, \rho^{hs})$  : tous les déplacements autour de  $\bar{X}$ , dans  $\mathbb{R}^2$ , de norme euclidienne  $\leq \rho^{hs}$  sont pris en considération ;
- 2i) pondérer, selon une fonction gaussienne, chaque terme de la fonction d'auto-dissemblance (équation (3.13)) en fonction de la distance à  $\bar{X}$  du point de  $\mathcal{Z}(\bar{X})$  ;
- 3i) faire précéder la recherche des maxima locaux d'un filtrage privilégiant les fonctions d'auto-dissemblance avec profil de cône circulaire très « pointu », ayant  $O$  pour sommet.

Il est à noter que le détecteur de Harris et Stephens est aussi connu sous le nom d'opérateur de Plessey.

## B) Les équations

1) La fonction d'auto-dissemblance  $\mathcal{D}^{I|\mathcal{Z}(\bar{X})}$  de Harris et Stephens est :

$$\begin{cases} \mathcal{D}^{I|\mathcal{Z}(\bar{X})} : B_2^{(f)}(O, \rho^{hs}) & \rightarrow \mathbb{R} \\ \phi = [\varphi \psi]^T & \mapsto \mathcal{D}^{I|\mathcal{Z}(\bar{X})}(\phi) \end{cases} \quad (3.43)$$

avec :

$$\begin{aligned} \mathcal{D}^{I|\mathcal{Z}(\bar{X})}(\phi) &= \sum_{X \in \mathcal{Z}(\bar{X})} G_{\sigma_I^2 I d_2}(\bar{X} - X) (I(X + \phi) - I(X))^2 \\ &= \sum_{X \in \mathcal{Z}(\bar{X})} \omega(X) \tilde{r}(X)^2, \end{aligned} \quad (3.44)$$

où :

- si  $C$  est une matrice symétrique (semi)-définie positive, dite de « variances-covariances », alors l'application gaussienne  $G_C$  en dimension 2 est définie par :

$$\begin{cases} G_C : \mathbb{R}^2 & \rightarrow \mathbb{R} \\ X = [x \ y]^T & \mapsto G_C(X) = \frac{1}{2\pi \sqrt{\text{Det}(C)}} e^{-X^T(C)^{-1}X}. \end{cases} \quad (3.45)$$

-  $\omega(X) = G_{\sigma_I^2 I d_2}(\bar{X} - X)$  est une pondération gaussienne centrée, d'écart type dit d'intégration  $\sigma_I$  ;

-  $\tilde{r}(X) = I(X + \phi) - I(X)$

2) La célèbre matrice  $M^{hs}$  : Harris et Stephens passent à une valeur approchée en conservant la partie régulière du développement de Taylor-Young à l'ordre 1 de l'expression de  $\tilde{r}(X)$  :

$$\tilde{r}(X) \approx (\nabla I(X))^T \phi \quad (3.46)$$

et

$$\begin{aligned} \mathcal{D}^{I|\mathcal{Z}(\bar{X})}(\phi) &\approx \tilde{\mathcal{D}}^{I|\mathcal{Z}(\bar{X})}(\phi) \\ &= \sum_{X \in \mathcal{Z}(\bar{X})} \omega(X) \phi^T \nabla I(X) (\nabla I(X))^T \phi \\ &= \phi^T \left( \sum_{X \in \mathcal{Z}(\bar{X})} \omega(X) \nabla I(X) (\nabla I(X))^T \right) \phi \\ &= \phi^T M^{hs}(\bar{X}, \sigma_I) \phi \end{aligned} \quad (3.47)$$

Cette matrice  $M^{hs}$  s'avère capitale et nous la retrouvons dans la plupart des détecteurs de points d'intérêt postérieurs à 1988.

$$M^{hs}(\bar{X}, \sigma_I) = \sum_{X \in \mathcal{Z}(\bar{X})} G_{\sigma_I^2 Id_2}(\bar{X} - X) \nabla I(X) (\nabla I(X))^T \in \mathcal{M}_2(\mathbb{R}). \quad (3.48)$$

Remarquons que :

$$\nabla I(X) (\nabla I(X))^T = \begin{bmatrix} \left(\frac{\partial I}{\partial x}(X)\right)^2 & \frac{\partial I}{\partial x}(X) \frac{\partial I}{\partial y}(X) \\ \frac{\partial I}{\partial x}(X) \frac{\partial I}{\partial y}(X) & \left(\frac{\partial I}{\partial y}(X)\right)^2 \end{bmatrix} \quad (3.49)$$

3) Montrons que la matrice  $M^{hs}$  recherchée doit satisfaire à :

$$0 < \tau_\lambda \leq \lambda_{\min}(M^{hs}) \approx \lambda_{\max}(M^{hs}), \quad (3.50)$$

où  $\lambda_{\min}(M^{hs})$  et  $\lambda_{\max}(M^{hs})$  représentent respectivement la plus petite et la plus grande valeur propre de  $M^{hs}$  et  $\tau_\lambda$  est un seuil bien choisi.

$M^{hs}$  est une matrice symétrique semi-définie positive car elle peut se mettre sous la forme :

$$M^{hs} = (J^{hs})^T J^{hs} \quad (3.51)$$

où :

$$J^{hs} = \begin{bmatrix} \sqrt{\omega(X_1)} \frac{\partial I}{\partial x}(X_1) & \sqrt{\omega(X_1)} \frac{\partial I}{\partial y}(X_1) \\ \vdots & \vdots \\ \sqrt{\omega(X_{n_{hs}})} \frac{\partial I}{\partial x}(X_{n_{hs}}) & \sqrt{\omega(X_{n_{hs}})} \frac{\partial I}{\partial y}(X_{n_{hs}}) \end{bmatrix} \in \mathcal{M}_{n_{hs}, 2}(\mathbb{R}) \quad (3.52)$$

(expression dans laquelle les points  $X$  de  $\mathcal{Z}(\bar{X})$  ont été ordonnés et indexés de 1 à  $n_{hs}$ ). On a donc  $\lambda_{\max}(M^{hs}) \geq \lambda_{\min}(M^{hs}) \geq 0$ .

De plus, rappelons que si  $\lambda_{\min}(M^{hs}) > 0$ , alors  $M^{hs}$  est définie positive.

$\forall \rho \in \mathbb{R}^{+*}$ , les problèmes :

$$(\mathcal{P}^{(min)}) \left\{ \min_{\|\phi\|_2 = \rho} (M^{hs} \phi | \phi) \right\} \text{ et } (\mathcal{P}^{(max)}) \left\{ \max_{\|\phi\|_2 = \rho} (M^{hs} \phi | \phi) \right\}$$

ont respectivement [[HU98], exo. 3.4, p. 85] :

i)  $\lambda_{\min}(M^{hs})\rho^2$  et  $\lambda_{\max}(M^{hs})\rho^2$  pour valeurs optimales.

2i)  $\phi^{\lambda_{\min}}$  (respectivement  $\phi^{\lambda_{\max}}$ ), vecteur propre associé à la valeur propre  $\lambda_{\min}(M^{hs})$  (respectivement  $\lambda_{\max}(M^{hs})$ ) pour solutions optimales.

Un point  $\bar{X}$  conduisant à  $\lambda_{\min}(M^{hs}) = \lambda_{\max}(M^{hs})$  est associé à une fonction d'auto-dissemblance pour laquelle, en notant  $\phi = [\varphi \psi]^T$ , la surface  $\Sigma$  :

$$\Sigma = \left\{ (\varphi, \psi, \mathcal{D}^{I|_{\mathcal{Z}(\bar{X})}}(\phi)) \text{ tels que } \phi \in B_2^{(f)}(O, \rho^{hs}) \right\}$$

est, en approximation du premier ordre, un tronc de cône circulaire d'axe « vertical », de sommet  $O$  et d'angle au sommet  $\alpha = \arctan(\frac{1}{\lambda_{\min}(M^{hs})})$ . Ce cône est d'autant plus pointu que  $\lambda_{\min}(M^{hs})$  est grand.

Ceci explique pourquoi on ne retient que les points  $\bar{X}$  et donc les matrices  $M^{hs}$  satisfaisant à  $0 < \tau_\lambda \leq \lambda_{\min}(M^{hs}) \approx \lambda_{\max}(M^{hs})$ .

---

### Remarque

Harris et Stephens justifient l'étude des valeurs propres en invoquant les courbures principales (voir Annexe C).

---

### 3.2.4.2 La pratique

#### A) Le critère de Harris et Stephens

Harris et Stephens [[HS88], p. 150] proposent donc le critère  $\mathcal{HS}$  suivant :

$$\begin{aligned}\mathcal{HS}(\bar{X}) &= \text{Det}(M^{hs}) - k \left( \text{Tr}(M^{hs}) \right)^2 \\ &= \lambda_{\min}(M^{hs}) \lambda_{\max}(M^{hs}) - k \left( \lambda_{\min}(M^{hs}) + \lambda_{\max}(M^{hs}) \right)^2\end{aligned}\quad (3.53)$$

Il est illustré sur la figure 3.6-(a).

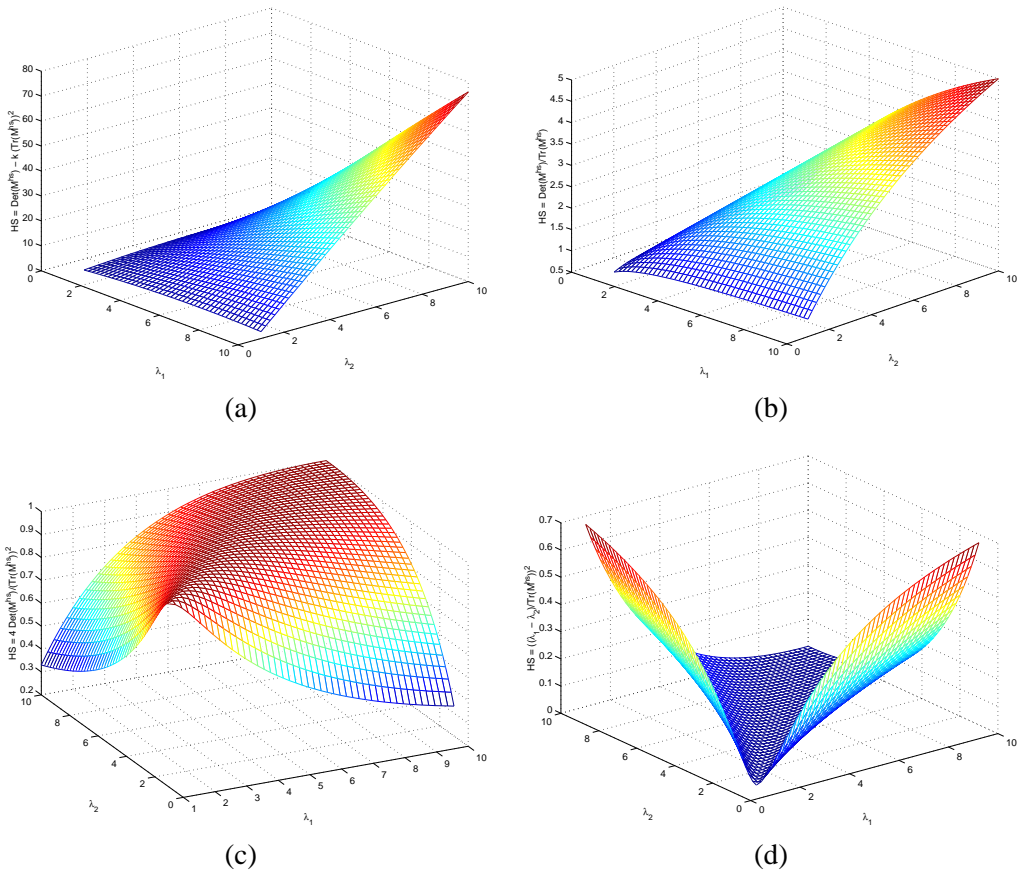


FIG. 3.6 – Tracés des critères de Harris et de ses variantes ( $\lambda_1 \stackrel{n}{=} \lambda_{\min}(M^{hs})$  et  $\lambda_2 \stackrel{n}{=} \lambda_{\max}(M^{hs})$ ) : (a) critère de Harris avec  $k = 0.04$  ; (b) variante de Haralick et Shapiro [[HS93b], éq. (13.74), p. 340] ; (c) variante de Haralick et Shapiro [[HS93b], éq. (13.75), p. 340] ; (d) variante de Urban [Urb03].

L'intérêt majeur de ce critère est d'éviter le calcul explicite des valeurs propres. À aucun moment, Harris et Stephens ne mentionnent une quelconque valeur stratégique pour  $k$ . Certaines implémentations utilisent  $k = 0.04$ . Cette valeur a été confirmée par Harris mais il ne sait pas comment elle a été diffusée. Elle est partie intégrante de l'implémentation « temps-réel » exploitant le jeu d'instructions MMX d'INTEL.

**B) Quels  $\rho^{(s)}$  pour définir  $\mathcal{Z}(\bar{X})$  ?**

À aucun moment, Harris et Stephens [HS88] ne mentionnent une quelconque valeur stratégique pour  $\rho^{(s)}$ . Interrogé, Harris nous a répondu que l'implémentation MMX utilise  $2\rho^{(s)} = 9$  associé à un noyau gaussien 2D à variables séparables.

**Relation entre binôme de Newton et loi normale** La probabilité de tirer  $k$  fois une boule d'une urne contenant autant de boules  $N$  (noires) que  $B$  (blanches) ( $p = q = 0.5$ ) au bout de  $n$  tirages avec remise est :

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} (0.5)^n. \quad (3.54)$$

Or cette loi de probabilité tend, pour  $n \rightarrow \infty$ , vers la loi normale  $\mathcal{N}(\mu, \sigma^2)$  où  $\mu = np$  et  $\sigma^2 = npq$ . Soit, ici,  $\mu = 0.5n$  et  $\sigma^2 = 0.25n$ . Ainsi, pour  $n = 3$ , on retrouve  $\mu = 1.5$ ,  $\sigma^2 = 0.75$  ( $\sigma = \sqrt{0.75}$ ) et, au coefficient multiplicatif  $(0.5)^3 (= 0.125)$  près, le noyau de convolution :

$$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \quad (3.55)$$

Pour  $n = 5$ , on retrouve  $\mu = 2.5$ ,  $\sigma^2 = 1.25$  et le noyau de convolution :

$$\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (3.56)$$

Le noyau 2D, à variables séparables,  $(9 \times 9)$  :

$$\begin{bmatrix} 1 & 2 & 8 & 14 & 16 & 14 & 7 & 2 & 1 \end{bmatrix} \quad (3.57)$$

utilisé par Harris dans son implémentation personnelle provient de :

$$\begin{bmatrix} 1 & 8 & 28 & 56 & 70 & 56 & 28 & 8 & 1 \end{bmatrix} \quad (3.58)$$

avec un facteur  $\frac{1}{4}$  et un ajustement aux extrémités (cependant  $70/4 = 17.5 \neq 16$ ). Il correspond à  $n = 9$ ,  $\mu = 4.5$  et  $\sigma^2 = 2.25$  ( $\sigma = 1.5$ ).

**C) Quels opérateurs de gradients ?**

Il s'agit tout simplement des noyaux de convolution 1D :

$$\frac{\partial I}{\partial x}(\bar{X}) \leftrightarrow \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \leftrightarrow \left( \frac{\partial I}{\partial y}(\bar{X}) \right)^T$$

provenant des polynômes orthogonaux discrets de Tchebychev à une indéterminée (à un facteur 2 près) et ne tiennent donc aucun compte de la nature bidimensionnelle des images.

**D) Quels  $\rho^{(m)}$  ?**

Les maxima locaux de la fonction  $\mathcal{HS}$  sont recherchés uniquement sur  $B_\infty^{(d)}(\bar{X}, 1)$ , d'où  $\rho^{(m)} = 1$  dans l'implémentation MMX.

### 3.2.4.3 Remarques

#### A) Des variantes pour $\mathcal{HS}$

- i) Haralick et Shapiro [[HS93b], éq. (13.74), p. 340] proposent une première variante, à un coefficient multiplicatif près :

$$\widetilde{\mathcal{HS}} = \frac{\text{Det}(M^{hs})}{\text{Tr}(M^{hs})}. \quad (3.59)$$

- 2i) Haralick et Shapiro [[HS93b], éq. (13.75), p. 340] proposent une seconde variante :

$$\widetilde{\mathcal{HS}} = 1 - \left( \frac{\lambda_{\min}(M^{hs}) - \lambda_{\max}(M^{hs})}{\lambda_{\min}(M^{hs}) + \lambda_{\max}(M^{hs})} \right)^2 = \frac{4 \text{Det}(M^{hs})}{(\text{Tr}(M^{hs}))^2}. \quad (3.60)$$

- 3i) Urban [Urb03] évoque une variante qui oblige au calcul des valeurs propres :

$$\widetilde{\mathcal{HS}} = \frac{\lambda_{\max}(M^{hs}) - \lambda_{\min}(M^{hs})}{\text{Tr}(M^{hs})}. \quad (3.61)$$

Ces variantes sont respectivement illustrées sur les figures 3.6-(b), (c) et (d).

#### B) Variante de Schmid *et al.*

Dans une étude comparative expérimentale relativement exhaustive [SMB00], Schmid *et al.* qualifient de version **améliorée** une implémentation  $\mathcal{HS}^{smb}$  de Harris et Stephens pour laquelle [[SMB00], §2.2, p. 156–157] :

- i)  $\nabla I$  est calculé à l'aide d'un noyau de convolution de  $G_{\sigma_D^2 Id_2}$ , où  $G_{\sigma_D^2 Id_2}$  est la fonction de Gauss centrée réduite, dont l'écart type  $\sigma_D$ , dit de « différentiation », vaut 1. On peut regretter l'absence de spécifications concernant la taille du noyau et son caractère 1D ou 2D.
- 2i)  $k = 0.06$
- 3i)  $\sigma_I = 2$
- 4i) le seuillage qui suit est mis en œuvre :

$$\forall X \in \mathbb{N}_n \times \mathbb{N}_m, \mathcal{HS}^{smb}(X) = \begin{cases} \mathcal{HS}(X) & \text{si } \mathcal{HS}(X) \geq \frac{1}{100} \left\{ \max_{\tilde{X} \in \mathbb{N}_n \times \mathbb{N}_m} \mathcal{HS}(\tilde{X}) \right\} \\ 0 & \text{sinon} \end{cases} \quad (3.62)$$

Étant donné les améliorations, la matrice  $M^{hs}$  intervenant dans l'évaluation de  $\mathcal{HS}$  devient la matrice  $M^{hs+smb}$ , où :

$$M^{hs+smb}(\bar{X}, \sigma_I, \sigma_D) = \sigma_D^2 \sum_{X \in \mathcal{Z}(\bar{X})} \nabla \tilde{I}_{\sigma_D^2 Id_2}(X) \left( \nabla \tilde{I}_{\sigma_D^2 Id_2}(X) \right)^T G_{\sigma_I^2 Id_2}(\bar{X} - X) \quad (3.63)$$

avec  $\tilde{I}_{\sigma_D^2 Id_2} = G_{\sigma_D^2 Id_2} \otimes I$ .

En généralisant cette matrice au cas où les matrices de variances-covariances  $C_I$  et  $C_D$  sont quelconques, on a :

$$M^{hs+smb}(\bar{X}, C_I, C_D) = \text{Det}(C_D) \sum_{X \in \mathcal{Z}(\bar{X})} \nabla \tilde{I}_{C_D}(X) \left( \nabla \tilde{I}_{C_D}(X) \right)^T G_{C_I}(\bar{X} - X) \quad (3.64)$$

où  $\tilde{I}_{C_D} = G_{C_D} \otimes I$ .

### 3.2.5 Détecteur de Shi et Tomasi : KLT

#### 3.2.5.1 La théorie

##### A) Le principe informel

L'objectif de Tomasi et Kanade [TK91] en 1991 puis de Shi et Tomasi [ST94] en 1994 est de caractériser les marques – points d'intérêt  $\bar{X}^{(i_0)}$  – de régions  $\mathcal{Z}^{(i_0)}(\bar{X}^{(i_0)})$  qui s'avéreront optimales lors de leur suivi d'une image  $I^{(i_0)}$  à l'image  $I^{(i_0+j_0)}$  d'une séquence vidéo  $\{I^{(0)}, \dots, I^{(i_0)}, \dots, I^{(i_0+j_0)}, \dots\}$ .

Le caractère d'optimalité doit être compris au sens où de telles régions  $\mathcal{Z}^{(i_0)}(\bar{X}^{(i_0)})$  permettent de calculer les paramètres de la transformation géométrique – translation [[TK91], éq. (2.2), p. 3] ou application affine [[ST94], éq. (2), p. 594] – qui déforme  $\mathcal{Z}^{(i_0)}(\bar{X}^{(i_0)})$  en  $\mathcal{Z}^{(i_0+j_0)}(\bar{X}^{(i_0+j_0)})$ .

Pour cela, ils utilisent tout simplement une mesure de corrélation – plus exactement de dissemblance – entre  $\mathcal{Z}^{(i_0)}(\bar{X}^{(i_0)})$  et  $\mathcal{Z}^{(i_0+j_0)}(\bar{X}^{(i_0+j_0)})$ .

Autrement dit, on remplace dans l'approche de Harris et Stephens [HS88] « auto-dissemblance » par « dissemblance » étant donné que l'image  $I^{(i_0+j_0)}$  vient jouer le rôle de  $I|_{\mathcal{Z}(\bar{X}+\phi)}$ .

##### B) Les équations

##### B.1) La mesure de dissemblance

Oublions la formulation dans l'espace  $L_2(\mathcal{Z}(\bar{X}))$  – qui suppose l'accès à la version continue du signal image !!! – utilisée par Tomasi et Kanade [[TK91], éq. (2.3), p. 3] et Shi et Tomasi [[ST94], éq. (3), p. 594] et plaçons nous dans  $(\mathbb{R}^{n_{kst}}, \|\cdot\|_2)$ , où  $n_{kst}$  est le nombre de pixels de  $\mathcal{Z}(\bar{X})$ , (comme au §4.5 page 124 du travail pionnier de Lucas et Kanade [LK81]) pour y exprimer la fonction de dissemblance sous la forme du critère d'un problème aux moindres carrés non linéaires.

Considérons par exemple les images  $I^{(0)}$  et  $I^{(1)}$  de la séquence ( $i_0 = 0$  et  $j_0 = 1$ ) et cherchons une transformation affine :

$$\begin{cases} T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ x \mapsto T(x) = Lx + t \end{cases} \quad (3.65)$$

$$\text{où } L = \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix} \text{ et } t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

La fonction de dissemblance  $\mathcal{D}^{I|_{\mathcal{Z}^{(0)}(\bar{X}^{(0)})}}$  choisie par Lucas et Kanade [[LK81], p. 123] est :

$$\begin{cases} \mathcal{D}^{I|_{\mathcal{Z}^{(0)}(\bar{X}^{(0)})}} : \mathbb{R}^6 \rightarrow \mathbb{R} \\ \beta \mapsto \mathcal{D}^{I|_{\mathcal{Z}^{(0)}(\bar{X}^{(0)})}}(\beta) = \|\tilde{R}^{\bar{X}^{(0)}}(\beta)\|_{(\mathbb{R}^{n_{kst}}, \|\cdot\|_2)}^2 \end{cases} \quad (3.66)$$

où  $\beta = [l_{11} - 1 \quad l_{21} \quad l_{12} \quad l_{22} - 1 \quad t_1 \quad t_2]^T$  et :

$$\left\{ \begin{array}{l} \tilde{R}^{\bar{X}^{(0)}} : \mathbb{R}^6 \rightarrow \mathbb{R}^{n_{kst}} \\ \beta \mapsto \tilde{R}^{\bar{X}^{(0)}}(\beta) = \begin{bmatrix} \vdots \\ \tilde{r}_i^{\bar{X}^{(0)}}(\beta) \\ \vdots \end{bmatrix}_{i=1..n_{kst}} \\ = \begin{bmatrix} \vdots \\ I^{(1)}(T(X_i^{(0)})) - I^{(0)}(X_i^{(0)}) \\ \vdots \end{bmatrix}_{i=1..n_{kst}} \end{array} \right. \quad (3.67)$$

avec  $\forall i = 1..n_{kst}, X_i^{(0)} \in \mathcal{Z}^{(0)}(\bar{X}^{(0)})$ .

### B.2) Le problème aux moindres carrés $(\mathcal{P}_{nls}^{\bar{X}^{(0)}})$ non linéaires

Le problème  $(\mathcal{P}_{nls}^{\bar{X}^{(0)}})$  à résoudre est le suivant :

$$(\mathcal{P}_{nls}^{\bar{X}^{(0)}}) \quad \begin{cases} \min \\ \beta \in \mathbb{R}^6 \end{cases} \mathcal{D}^{I|_{\mathcal{Z}^{(0)}(\bar{X}^{(0)})}}(\beta) \quad (3.68)$$

Notons que ce critère suppose la conservation des niveaux de gris des couples de points homologues, à savoir :

$$I^{(0)}(X_i^{(0)}) \approx I^{(1)}(T(X_i^{(0)})), \forall i = 1..n_{kst}. \quad (3.69)$$

### B.3) Passage du problème $(\mathcal{P}_{nls}^{\bar{X}^{(0)}})$ à un problème aux moindres carrés linéaires $(\mathcal{P}_{ls}^{\bar{X}^{(0)}})$

Posons  $\forall i = 1..n_{kst}, X_i^{(0)} + \phi_i^{(0)} = T(X_i^{(0)})$  et supposons  $\phi_i^{(0)}$  « petit » pour approcher  $I^{(1)}(T(X_i^{(0)}))$  par la partie régulière du développement de Taylor-Young à l'ordre 1 :

$$\begin{aligned} I^{(1)}(T(X_i^{(0)})) &= I^{(1)}(X_i^{(0)} + \phi_i^{(0)}) \\ &\approx I^{(1)}(X_i^{(0)}) + \nabla I^{(1)}(X_i^{(0)})^T \phi_i^{(0)} \\ &\approx I^{(1)}(X_i^{(0)}) + \nabla I^{(1)}(X_i^{(0)})^T [(L - \text{Id}_2)X_i^{(0)} + t] \end{aligned} \quad (3.70)$$

En notant  $I_x^{(1)} = \frac{\partial I^{(1)}}{\partial x}(X_i^{(0)})$  et  $I_y^{(1)} = \frac{\partial I^{(1)}}{\partial y}(X_i^{(0)})$ , alors l'équation (3.70) s'écrit :

$$I^{(1)}(X_i^{(0)} + \phi_i^{(0)}) \approx [I_x^{(1)} x_i^0 \quad I_y^{(1)} x_i^0 \quad I_x^{(1)} y_i^0 \quad I_y^{(1)} y_i^0 \quad I_x^{(1)} \quad I_y^{(1)}] \beta + I^{(1)}(X_i^{(0)}) \quad (3.71)$$

d'où :

$$\forall i = 1..n_{kst}, \tilde{r}_i^{\bar{X}^{(0)}}(\beta) = C(X_i^{(0)})\beta + I^{(1)}(X_i^{(0)}) - I^{(0)}(X_i^{(0)}), \quad (3.72)$$

avec :

$$C(X_i^{(0)}) = [I_x^{(1)} x_i^0 \quad I_y^{(1)} x_i^0 \quad I_x^{(1)} y_i^0 \quad I_y^{(1)} y_i^0 \quad I_x^{(1)} \quad I_y^{(1)}]. \quad (3.73)$$



En reportant l'équation précédente dans le problème  $(\mathcal{P}_{nls}^{\bar{X}^{(0)}})$ , on aboutit au problème aux moindres carrés linéaires  $(\mathcal{P}_{ls}^{\bar{X}^{(0)}})$  suivant :

$$(\mathcal{P}_{ls}^{\bar{X}^{(0)}}) \quad \begin{cases} \min & \|A\beta - b\|_{(\mathbb{R}^{n_{kst}}, \|\cdot\|_2)}^2 \\ \beta \in \mathbb{R}^6 \end{cases} \quad (3.74)$$

où :

$$\begin{aligned} A &= [A_1 \quad A_2] \\ A_1 &= \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ I_x^{(1)} x_i^0 & I_y^{(1)} x_i^0 & I_x^{(1)} y_i^0 & I_y^{(1)} y_i^0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}_{i=1..n_{kst}} \\ A_2 &= \begin{bmatrix} \vdots & \vdots \\ I_x^{(1)} & I_y^{(1)} \\ \vdots & \vdots \end{bmatrix}_{i=1..n_{kst}} \\ b &= \begin{bmatrix} \vdots \\ -I^{(1)}(X_i^{(0)}) + I^{(0)}(X_i^{(0)}) \\ \vdots \end{bmatrix}_{i=1..n_{kst}} \end{aligned} \quad (3.75)$$

#### B.4) Résolution du problème aux moindres carrés linéaires $(\mathcal{P}_{ls}^{\bar{X}^{(0)}})$

$\hat{\beta}$  satisfait aux équations normales de Gauss [[Bjö96], éq. (9.2.2), p. 343] :

$$(A^T A) \hat{\beta} = A^T b \quad (3.76)$$

avec  $A^T A = \begin{bmatrix} A_1^T A_1 & A_1^T A_2 \\ A_2^T A_1 & A_2^T A_2 \end{bmatrix} = \begin{bmatrix} U & V \\ V^T & W \end{bmatrix}$ , où :

$$U = \sum_{i=1}^{n_{kst}} \begin{bmatrix} (I_x^{(1)})^2 (x_i^0)^2 & I_x^{(1)} I_y^{(1)} (x_i^0)^2 & (I_x^{(1)})^2 x_i^0 y_i^0 & I_x^{(1)} I_y^{(1)} x_i^0 y_i^0 \\ I_x^{(1)} I_y^{(1)} (x_i^0)^2 & (I_y^{(1)})^2 (x_i^0)^2 & I_x^{(1)} I_y^{(1)} x_i^0 y_i^0 & (I_y^{(1)})^2 x_i^0 y_i^0 \\ (I_x^{(1)})^2 x_i^0 y_i^0 & I_x^{(1)} I_y^{(1)} x_i^0 y_i^0 & (I_x^{(1)})^2 (y_i^0)^2 & I_x^{(1)} I_y^{(1)} (y_i^0)^2 \\ I_x^{(1)} I_y^{(1)} x_i^0 y_i^0 & (I_y^{(1)})^2 x_i^0 y_i^0 & I_x^{(1)} I_y^{(1)} (y_i^0)^2 & (I_y^{(1)})^2 (y_i^0)^2 \end{bmatrix} \quad (3.77)$$

$$V^T = \sum_{i=1}^{n_{kst}} \begin{bmatrix} (I_x^{(1)})^2 x_i^0 & I_x^{(1)} I_y^{(1)} x_i^0 & (I_x^{(1)})^2 y_i^0 & I_x^{(1)} I_y^{(1)} y_i^0 \\ I_x^{(1)} I_y^{(1)} x_i^0 & (I_y^{(1)})^2 x_i^0 & I_x^{(1)} I_y^{(1)} y_i^0 & (I_y^{(1)})^2 y_i^0 \end{bmatrix}^T \quad (3.78)$$

$$W = \sum_{i=1}^{n_{kst}} \begin{bmatrix} (I_x^{(1)})^2 & I_x^{(1)} I_y^{(1)} \\ I_x^{(1)} I_y^{(1)} & (I_y^{(1)})^2 \end{bmatrix} \quad (3.79)$$

On retrouve pour  $W$ , au facteur de pondération  $\omega$  près, la matrice  $M^{hs}$  (équation (3.48)).

### B.5) Cas particulier où $L = Id_2$

On a  $\beta_i = 0, \forall i = 1..4$  et la transformation  $T$  se ramène à une translation, ce qui est a priori un modèle acceptable dans le cas de deux images successives d'une séquence, rapprochées dans le temps ou correspondant à de faibles variations de la géométrie (scène  $\leftrightarrow$  caméra).

Les équations normales de Gauss (équation (3.76)) se ramènent à :

$$W \begin{bmatrix} \beta_5 \\ \beta_6 \end{bmatrix} = A_2^T b \quad (3.80)$$

La « qualité » de la solution de l'équation (3.80) dépend du nombre de conditionnement de  $W$ .  $W$  est une matrice symétrique semi-définie positive dans tous les cas. Naturellement, toute configuration qui aboutirait à une matrice singulière  $W$  doit être rejetée a priori de sorte que  $W$  est définie positive et par conséquent ses valeurs propres et valeurs singulières sont confondues et le nombre de conditionnement est  $Cond_2(W) = \frac{\lambda_{max}(W)}{\lambda_{min}(W)}$ .

Le cas le plus favorable étant  $\lambda_{max}(W) = \lambda_{min}(W)$  ( $Cond_2(W) = 1$  est alors optimal), Shi et Tomasi [ST94] proposent de rejeter les régions  $\mathcal{Z}(\bar{X}^{(0)})$  caractérisées par  $\lambda_{min} \leq \tau_\lambda$ , où  $\tau_\lambda$  est un seuil bien choisi.

---

#### Remarque

À l'instar de Harris et Stephens [HS88], Shi et Tomasi pondèrent la  $i$ -ème composante du vecteur  $A\beta - b$  par  $\sqrt{\omega(X_i^{(0)})}$ , induit par une fonction gaussienne.

---

### 3.2.5.2 La pratique

#### A) Le critère de Shi et Tomasi

Shi et Tomasi [[ST94], éq. (8), p. 595] proposent le critère suivant :

$$ST(\bar{X}) = \begin{cases} 0 & \text{si } \lambda_{min}(W) \leq \tau_\lambda \\ \lambda_{min}(W) & \text{sinon} \end{cases} \quad (3.81)$$

#### B) À propos de $\tau_\lambda$

Tomasi et Kanade [[TK91], p. 8] « étalonnent » préalablement la caméra en :

- i) déterminant  $\tau_\lambda^{min}$  correspondant à des valeurs propres  $\lambda_{min}(W)$  associées à des régions sensiblement uniformes ;
- 2i) déterminant  $\tau_\lambda^{max}$  correspondant à des valeurs propres  $\lambda_{min}(W)$  associées à des coins, des zones hautement texturées...
- 3i) choisissant  $\tau_\lambda = \frac{\tau_\lambda^{min} + \tau_\lambda^{max}}{2}$ .

#### C) À propos de la recherche des maxima locaux

Cette recherche est particulièrement simplifiée en contraignant une intersection vide entre deux régions quelconques (voir les illustrations de Tomasi et Kanade [[TK91], figures 5.5 et 5.8, p. 13 et 15] et de Shi et Tomasi [[ST94], figure 11, p. 598]). En contrepartie, cette stratégie raréfie les points d'intérêt.

**D) À propos de la taille des régions  $\mathcal{Z}(\bar{X}^{(0)})$** 

Il y a un compromis à définir automatiquement la taille des régions  $\mathcal{Z}(\bar{X}^{(0)})$  [OK90, Cha94], car les petites régions sont sensibles au bruit tandis que les grandes régions sont affectées par les occultations, etc. Dans les expériences conduites par Tomasi et Kanade, et Shi et Tomasi,  $2\rho^{(s)} = \{14, \dots, 24, 30\}$ .

**E) Le choix du modèle : transformation affine ou translation**

Ce choix est du ressort de l'étape de suivi et il faut bien être conscient que seule la matrice  $W$  importe pour le choix des points d'intérêt.

La différence des critères retenus par Harris et Stephens [HS88] et Tomasi, Shi et Kanade [TK91, ST94] s'explique par l'objectif d'exploitation des points d'intérêt dans une étape de suivi, pour ces derniers auteurs.

**F) À propos du logiciel du domaine public**

Mis à jour régulièrement par Birchfield, il est accessible sur le site : <http://www.ces.clemson.edu/~stb/klt/>.

L'utilisateur a la possibilité d'ajuster :

- i)  $\sigma$ , l'écart type définissant le noyau de convolution (gaussienne et sa dérivée)
- 2i)  $\tau_\lambda$
- 3i) la distance minimale séparant deux maxima locaux
- 4i)  $\rho^{(s)}$

**3.2.6 Détecteur de Smith et Brady : SUSAN****3.2.6.1 La théorie****A) Le principe informel**

L'opérateur SUSAN [SB95] proposé par Smith et Brady peut aussi bien détecter des points de contours que des coins ou des repères de zones hautement texturées. Nous présentons la version paramétrée pour détecter des coins.

Un coin traduit une discontinuité de la fonction image  $I$  dans  $d (\geq 2)$  directions. En conséquence, une première condition nécessaire (CN1) pour qu'un pixel  $\bar{X}$  référence un coin est que son niveau de gris diffère, de manière significative, de celui d'une majorité de ses voisins. L'ensemble des pixels voisins de niveau de gris semblable à  $I(\bar{X})$  est appelé USAN (*Univalue Segment Assimilation Nucleus*) et cette CN1 élit des points  $\bar{X}$  associés à des valeurs  $USAN(\bar{X})$  faibles (*Smallest USAN*). La figure 3.7 donne une illustration des zones USAN.

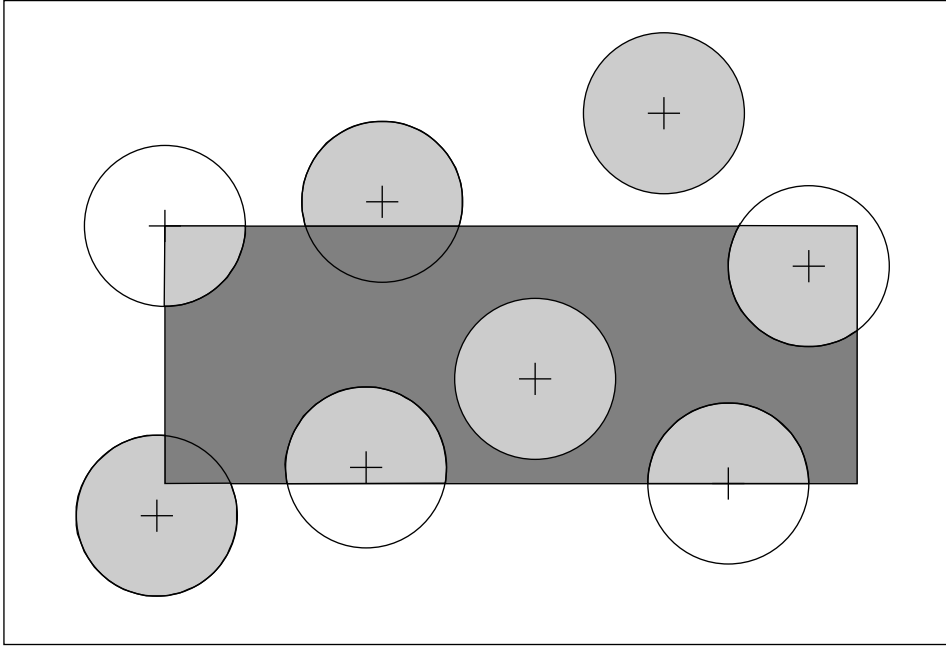


FIG. 3.7 – Image de synthèse composée d'un rectangle noir sur un fond blanc. Pour chacune des huit régions  $\mathcal{Z}(\bar{X})$ , où  $\bar{X}$  est symbolisé par le '+', la zone  $\text{USAN}(\bar{X})$  est représentée en grisé.

Une seconde condition nécessaire (CN2) pour que le pixel  $\bar{X}$  référence un coin est que la région  $\text{USAN}(\bar{X})$  soit d'un seul tenant ou, plus formellement, une composante connexe (par arcs) de  $\bar{X}$ .

## B) Les équations

### B.1) À propos de la (CN1) : calcul de la région $\text{USAN}(\bar{X})$

- a) L'outil est une mesure d'auto-corrélation (ressemblance) appliquée à la fonction  $I_{|\{\bar{X}\}}$ , à savoir [[SB97], éq. (4), p. 50 et figure 5, p. 47] :

$$\begin{cases} \mathcal{S}^{I_{|\{\bar{X}\}}}: B_2^{(d)}(\bar{X}, \rho^{(s)}) \rightarrow ]0, 1] \\ X \mapsto \mathcal{S}^{I_{|\{\bar{X}\}}}(X) = e^{-\left(\frac{I(X) - I(\bar{X})}{\tau_u}\right)^6} \end{cases} \quad (3.82)$$

où  $\tau_u$  est un seuil sur le contraste maximal autorisé dans  $\text{USAN}(\bar{X})$ . L'application  $\mathcal{S}^{I_{|\{\bar{X}\}}}$  est illustrée sur la figure 3.8.

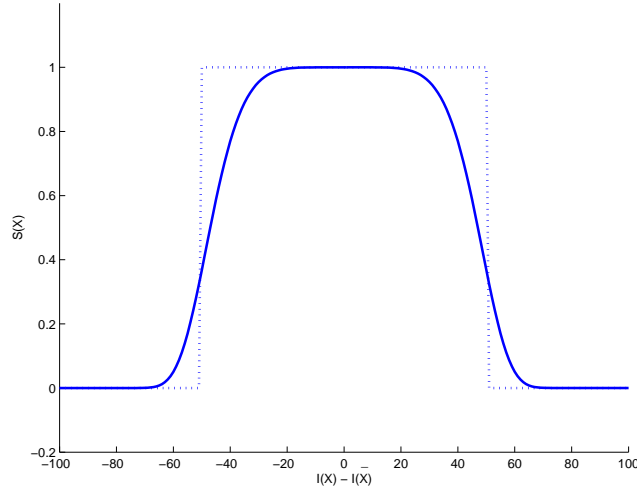


FIG. 3.8 – Tracé (en trait plein) de la mesure d'auto-correlation  $\mathcal{S}^{I_{\{\bar{X}\}}}$  (notée  $S$  sur ce graphique) en fonction de la différence de niveaux de gris pour  $\tau_u = 50$ . En trait discontinu, la fonction « créneau » sur  $[-\tau_u, \tau_u]$ .

- b) La fonction  $\mathcal{M}$  (équation (3.18)) de l'opérateur de Moravec est ici remplacée par [[SB97], éq. (3), p. 49] :

$$\begin{cases} \mathcal{SB} : \mathbb{N}_n \times \mathbb{N}_m \rightarrow \mathbb{R} \\ \bar{X} \mapsto \mathcal{SB}(\bar{X}) = \max \left( 0, \tau_g - \sum_{X \in B_2^{(d)}(\bar{X}, \rho^{(s)})} \mathcal{S}^{I_{\{\bar{X}\}}}(X) \right) \end{cases} \quad (3.83)$$

où  $\tau_g$  (éométrie) fixe la proportion maximale de points d'une région  $\text{USAN}(\bar{X})$  éligible.

### B.2) À propos de la (CN2)

- a) Le calcul du barycentre  $\bar{G}$  de la région  $\text{USAN}(\bar{X})$  est donné par :

$$\bar{G} = \frac{\sum_{X \in \text{USAN}(\bar{X})} \mathcal{S}^{I_{\{\bar{X}\}}}(X) X}{\sum_{X \in \text{USAN}(\bar{X})} \mathcal{S}^{I_{\{\bar{X}\}}}(X)} \quad (3.84)$$

- b) La caractérisation d'une région  $\text{USAN}(\bar{X})$  connexe est la suivante :

- i)  $\|\vec{\bar{X}\bar{G}}\| > \tau_c$
- 2i)  $\forall X \in B_2^{(d)}(\bar{X}, \rho^{(s)}) \cap \{\lambda \vec{\bar{X}\bar{G}} | \lambda \in \mathbb{R}^{+*}\} \Rightarrow X \in \text{USAN}(\bar{X})$

### B.3) Quels $\bar{X}$ ?

$\bar{X}$  est retenu point d'intérêt au sens SUSAN si et seulement si :

$$\begin{cases} \bar{X} \text{ satisfait à la CN2} \\ \bar{X} = \left\{ \begin{array}{l} \text{argmax} \\ X \in B(\bar{X}, \rho^{(m)}) \end{array} \mathcal{SB}(X) \right\} \end{cases} \quad (3.85)$$

### 3.2.6.2 La pratique

- i)  $\rho^{(s)} = 3$  en général (36 voisins de  $\bar{X}$ ) et 1.5 dans les cas d'école [[SB97], §4.2, p. 49]
- 2i)  $\tau_u \in \{7, 10, 15, 25, 27\}$
- 3i)  $\tau_g \leq 0.5\pi(\rho^{(s)})^2$  [[SB97], §5.2, p. 60]
- 4i)  $\rho^{(m)} = 2.5$  [[SB97], §5.2, p. 60]
- 5i)  $\tau_c$  : Smith et Brady ne fournissent aucune indication mais nos expériences personnelles utilisent  $\tau_c > \frac{\rho^{(s)}}{3}$

### 3.2.7 Détecteur de Trajkovic et Hedley

#### 3.2.7.1 La théorie

##### A) Le principe informel

L'objectif du détecteur de Trajkovic et Hedley [TH98] est de détecter des coins, toutes catégories confondues : jonctions « L », « X », « Y », « T »... (voir figure 3.1).

Lorsque  $\bar{X}$  marque un coin, alors, par construction géométrique, pour toute boule  $B_2^{(d)}(\bar{X}, \rho^{(s)})$  (disque), tout diamètre de celle-ci rencontre la sphère  $S_2^{(d)}(\bar{X}, \rho^{(s)})$  en deux points  $m(\bar{X})$  et  $m'(\bar{X})$ , l'un au moins (disons  $m(\bar{X})$ ) n'appartenant pas à  $\text{USAN}(\bar{X})$  (voir sur la figure 3.9, le diamètre en tirets bleus).

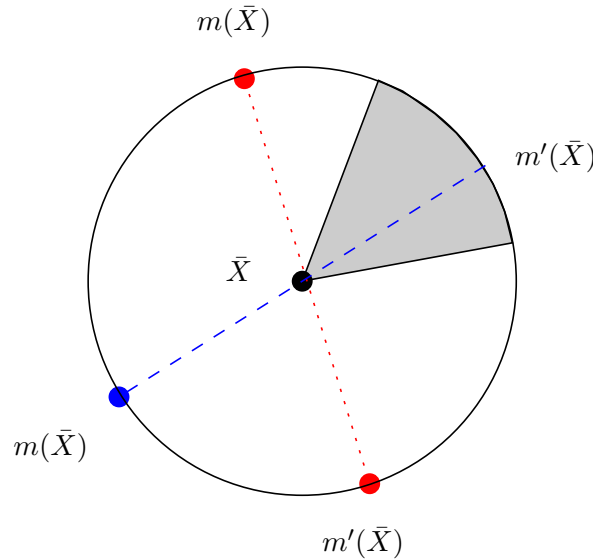


FIG. 3.9 – Illustration du principe de Trajkovic et Hedley. En grisé, la zone  $\text{USAN}(\bar{X})$ . Deux diamètres sont représentés : en pointillé rouge, ni  $m(\bar{X})$  ni  $m'(\bar{X})$  n'appartiennent à la zone  $\text{USAN}(\bar{X})$ ; en tiret bleu,  $m(\bar{X})$  n'appartient pas à la zone  $\text{USAN}(\bar{X})$  donc la propriété est vérifiée.

En conséquence, la dissemblance locale de  $I_{\{\bar{X}\}}$  (voir Moravec, §3.2.2, équation (3.13) et Harris, §3.2.4, équation (3.43)) pour des déplacements  $\phi$  amenant  $\bar{X}$  sur la sphère  $S_2^{(d)}(\bar{X}, \rho^{(s)})$  est nécessairement toujours forte.

## B) Les équations

### B.1) La fonction de dissemblance

Elle est définie par :

$$\left\{ \begin{array}{ll} \mathcal{D}^{I_{|\{\bar{X}\}}}: S_2^{(d)}(O, \rho^{(s)}) & \rightarrow \mathbb{R} \\ \phi & \mapsto \mathcal{D}^{I_{|\{\bar{X}\}}}(\phi) = (I(\bar{X}) - I(\bar{X} + \phi))^2 \\ & \quad + (I(\bar{X}) - I(\bar{X} - \phi))^2 \end{array} \right. \quad (3.86)$$

Notons qu'avec les notations de la figure 3.9,  $m(\bar{X}) = \bar{X} + \phi$  et  $m'(\bar{X}) = \bar{X} - \phi$ .

### B.2) Le critère de Trajkovic et Hedley

Il est donné par [[TH98], éq. (3), p. 78] :

$$\left\{ \begin{array}{ll} \mathcal{TH}: \mathbb{N}_n \times \mathbb{N}_m & \rightarrow (\mathbb{N} \subset) \mathbb{R} \\ \bar{X} & \mapsto \mathcal{TH}(\bar{X}) = \left\{ \min_{\phi \in S_2^{(d)}(O, \rho^{(s)})} \mathcal{D}^{I_{|\{\bar{X}\}}}(\phi) \right\} \end{array} \right. \quad (3.87)$$

### B.3) Le calcul de $\mathcal{TH}(\bar{X})$

- a) Trajkovic et Hedley [TH98] proposent un calcul approché s'appuyant uniquement sur :

$$\left\{ I(\bar{X}), \underbrace{I(\bar{X} + \overbrace{(\rho^{(s)}, 0)}^{\phi^{(1)}})}_{\bar{X}^{(1)}}, \underbrace{I(\bar{X} + \overbrace{(0, \rho^{(s)})}^{\phi^{(2)}})}_{\bar{X}^{(2)}}, \underbrace{I(\bar{X} + \overbrace{(-\rho^{(s)}, 0)}^{\phi^{(3)}})}_{\bar{X}^{(3)}}, \underbrace{I(\bar{X} + \overbrace{(0, -\rho^{(s)})}^{\phi^{(4)}})}_{\bar{X}^{(4)}} \right\}$$

et  $\{\mathcal{D}^{I_{|\{\bar{X}\}}}(\phi^{(i)})\}_{i=1,\dots,4}$  associé.

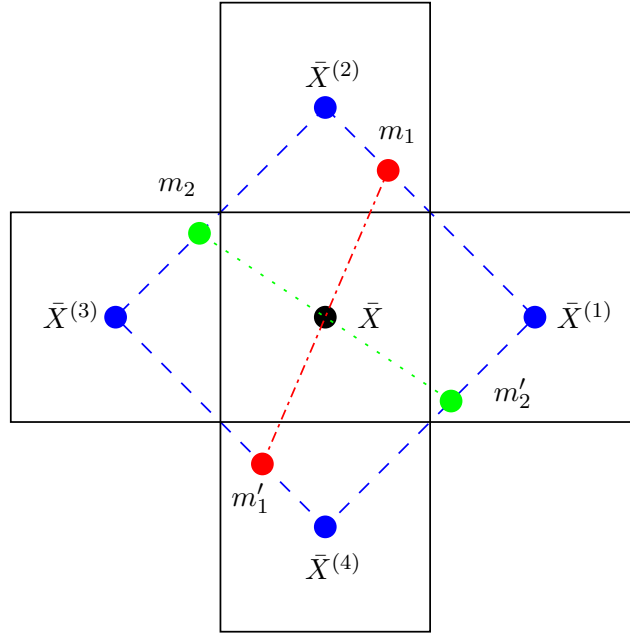
- b) Pour un déplacement  $\phi \in S_1(O, \rho^{(s)})$  associé aux points diamétralement opposés  $m$  et  $m'$ , Trajkovic et Hedley proposent une valeur approchée de  $I(m)$  et  $I(m')$  obtenue en considérant une variation linéaire entre respectivement :

$$\begin{cases} I(\bar{X}^{(1)}) \text{ et } I(\bar{X}^{(2)}) \\ I(\bar{X}^{(3)}) \text{ et } I(\bar{X}^{(4)}) \end{cases} \quad \text{si } m \in [\bar{X}^{(1)}, \bar{X}^{(2)}] \text{ (donc si } m' \in [\bar{X}^{(3)}, \bar{X}^{(4)}])$$

ou :

$$\begin{cases} I(\bar{X}^{(2)}) \text{ et } I(\bar{X}^{(3)}) \\ I(\bar{X}^{(4)}) \text{ et } I(\bar{X}^{(1)}) \end{cases} \quad \text{si } m \in [\bar{X}^{(2)}, \bar{X}^{(3)}] \text{ (donc si } m' \in [\bar{X}^{(4)}, \bar{X}^{(1)}])$$

- c) Dans le cas où  $\rho^{(s)} = 1$ , une illustration est donnée sur la figure 3.10.

FIG. 3.10 – Illustration du principe de Trajkovic et Hedley pour  $S_1(O, 1)$ 

d) Les expressions analytiques sont les suivantes :

$$\forall i = 1, 2, \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(i)}) = (I(\bar{X}) - I(\bar{X}^{(i)}))^2 + (I(\bar{X}) - I(\bar{X}^{(i+2)}))^2 \quad (3.88)$$

et :

$$\forall \phi \in S_1(O, 1), \mathcal{D}^{I_{\{\bar{X}\}}}(\phi) = (I(\bar{X}) - I(m))^2 + (I(\bar{X}) - I(m'))^2, \quad (3.89)$$

où :

$$\begin{cases} I(m) = (1 - \alpha)I(\bar{X}^{(1)}) + \alpha I(\bar{X}^{(2)}) \\ I(m') = (1 - \alpha)I(\bar{X}^{(3)}) + \alpha I(\bar{X}^{(4)}) \end{cases} \quad \text{si } m \stackrel{n}{=} m_1 \in [\bar{X}^{(1)}, \bar{X}^{(2)}] \text{ et } \alpha \in [0, 1]$$

ou bien :

$$\begin{cases} I(m) = (1 - \alpha)I(\bar{X}^{(3)}) + \alpha I(\bar{X}^{(2)}) \\ I(m') = (1 - \alpha)I(\bar{X}^{(1)}) + \alpha I(\bar{X}^{(4)}) \end{cases} \quad \text{si } m \stackrel{n}{=} m_2 \in [\bar{X}^{(2)}, \bar{X}^{(3)}] \text{ et } \alpha \in [0, 1]$$

En reportant ces expressions dans celle de  $\mathcal{D}^{I_{\{\bar{X}\}}}(\phi)$  (équation (3.86)), on obtient aisément [[TH98], éq. (9), p. 78] :

$$\mathcal{D}^{I_{\{\bar{X}\}}}(\phi) = \begin{cases} a_1 \alpha^2 + 2b_1 \alpha + c, & \text{si } m = m_1 \\ a_2 \alpha^2 + 2b_2 \alpha + c, & \text{si } m = m_2 \end{cases} \quad (3.90)$$

où :

$$\begin{cases} c &= \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(1)}) \\ b_1 &= (I(\bar{X}^{(2)}) - I(\bar{X}^{(1)}))(I(\bar{X}^{(1)}) - I(\bar{X})) \\ &\quad + (I(\bar{X}^{(4)}) - I(\bar{X}^{(3)}))(I(\bar{X}^{(3)}) - I(\bar{X})) \\ b_2 &= (I(\bar{X}^{(2)}) - I(\bar{X}^{(3)}))(I(\bar{X}^{(3)}) - I(\bar{X})) \\ &\quad + (I(\bar{X}^{(4)}) - I(\bar{X}^{(1)}))(I(\bar{X}^{(1)}) - I(\bar{X})) \\ a_1 &= \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(2)}) - \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(1)}) - 2b_1 \\ a_2 &= \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(2)}) - \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(1)}) - 2b_2 \end{cases} \quad (3.91)$$



En posant  $\begin{cases} b = \min(b_1, b_2) \\ a = \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(2)}) - \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^{(1)}) - 2b \end{cases}$ , alors le critère de Trajkovic et Hedley  $\mathcal{TH}$  se réécrit :

$$\mathcal{TH}(\bar{X}) = \begin{cases} c - \frac{b^2}{a} & \text{si et seulement si } (b < 0 \text{ et } a + b > 0) \\ \min_{i=1,2} \mathcal{D}^{I_{\{\bar{X}\}}}(\phi^i) & \text{sinon (valeur de dépannage)} \end{cases} \quad (1)$$

(3.92)

e) Extension à la sphère  $S_2(O, \rho^{(s)})$  :

Il s'agit de considérer une variation linéaire de  $I(m)$  (respectivement  $I(m')$ ) non plus sur le segment de droite  $[\bar{X}^{(1)}, \bar{X}^{(2)}]$  (respectivement  $[\bar{X}^{(3)}, \bar{X}^{(4)}]$ ) si  $m \in [\bar{X}^{(1)}, \bar{X}^{(2)}]$  et  $[\bar{X}^{(2)}, \bar{X}^{(3)}]$  (respectivement  $[\bar{X}^{(4)}, \bar{X}^{(1)}]$ ) si  $m \in [\bar{X}^{(2)}, \bar{X}^{(3)}]$ , mais sur l'arc de cercle  $\widehat{[\bar{X}^{(1)}, \bar{X}^{(2)}]}$  (respectivement  $\widehat{[\bar{X}^{(3)}, \bar{X}^{(4)}]}$ ) si  $m \in \widehat{[\bar{X}^{(1)}, \bar{X}^{(2)}]}$  et  $\widehat{[\bar{X}^{(2)}, \bar{X}^{(3)}]}$  (respectivement  $\widehat{[\bar{X}^{(4)}, \bar{X}^{(1)}]}$ ) si  $m \in \widehat{[\bar{X}^{(2)}, \bar{X}^{(3)}]}$  (voir illustration sur la figure 3.11).

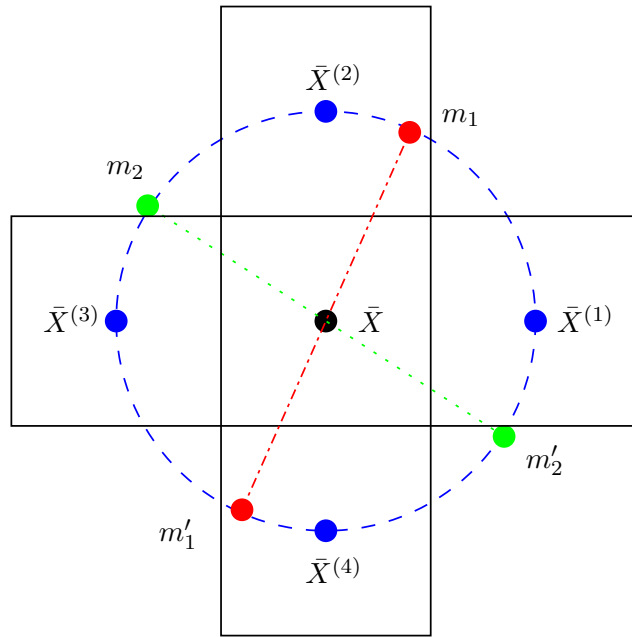


FIG. 3.11 – Illustration du principe de Trajkovic et Hedley pour  $S_2(O, 1)$

Une solution analytique existe pour  $\mathcal{TH}(\bar{X})$  car  $\mathcal{D}^{I_{\{\bar{X}\}}}(\phi)$  se met sous la forme [[TH98], éq. (14), p. 79] :

$$a' \cos(\theta) + b' \sin(\theta) + c \quad (3.93)$$

f) Extension à la sphère  $S_\infty(O, \rho^{(s)})$  ( $\rho^{(s)} = 1$  sur l'illustration de la figure 3.12) :

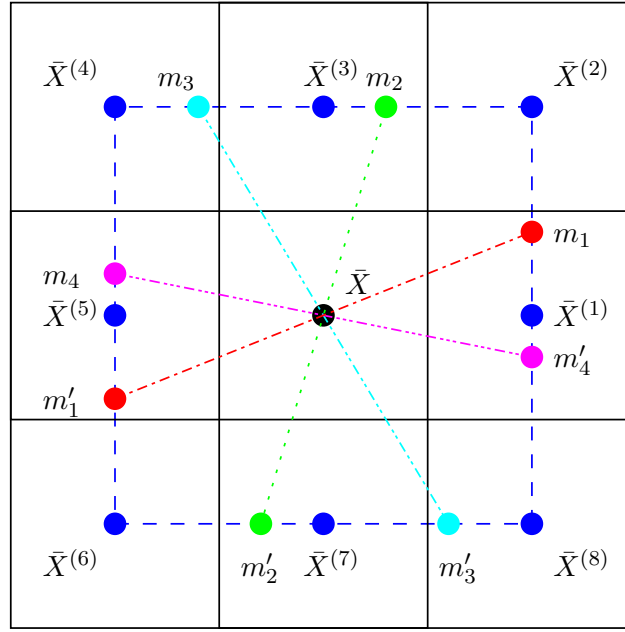


FIG. 3.12 – Illustration du principe de Trajkovic et Hedley pour  $S_\infty(O, 1)$

Il y a autant de points  $\bar{X}^{(i)}$  que de pixels appartenant à la sphère (8 pour  $\rho^{(s)} = 1$ ) et les interpolations linéaires sur les segments  $[\bar{X}^{(i)}, \bar{X}^{(i+1)}]$  ( $i = 1..4$  pour  $\rho^{(s)} = 1$ ) s'en déduisent.

### 3.2.7.2 La pratique

#### A) La multi-résolution

Trajkovic et Hedley appliquent successivement :

- i) l'équation (3.92)-(2) sur une version sous-échantillonnée de l'image de façon à éliminer par seuillage ( $\leq \tau_G^{(1)}$ ) les coins « texture » associés à des effets (de bord) de la texture par opposition aux coins « géométriques » associés aux contours délimitant les objets ( $\tau_G^{(1)} \in [0, 200]$  ( $\tau_G^{(1)} \approx 50$  est un bon compromis)) ;
- 2i) l'équation (3.92)-(2) sur les coins potentiels provenant de l'étape précédente (i) lors du passage à la pleine résolution de façon à éliminer par seuillage ( $\leq \tau_G^{(2)}$ ) les coins « texture » associés à des effets (de bord) de la texture par opposition aux coins « géométriques » associés aux contours délimitant les objets ( $\tau_G^{(2)} \in [200, 800]$  ( $\tau_G^{(2)} \approx 500$  est un bon compromis)) ;
- 3i) l'équation (3.92)-(1) aux pixels « rescapés » de l'étape précédente (2i) avec à nouveau un seuillage (le même  $\tau_G^{(2)}$ ).

#### B) Quels $\rho^{(s)}$ ?

Les rayons de la boule intervenant dans la fonction de dissemblance sont différents selon l'étape de l'algorithme [[TH98], §6.3, p. 81] :

- i) l'adaptation de l'équation (3.92) est appliquée pour  $S_2(O, \rho^{(s)})$ , où  $\rho^{(s)} = 3$  (soit  $\bar{X}^{(1)}, \dots, \bar{X}^{(16)}$  donc  $\phi^{(1)}, \dots, \phi^{(16)}$ ) à l'étape (2i) ;

- 2i) l'étape (3i) a été implémentée pour  $S_1(O, 1)$  (équation (3.92)-(1)) et pour  $S_2(O, 1)$  (voir l'extension présentée à la section 3.2.7.1-B.3.e).

### C) Quel $\rho^{(m)}$ ?

Le rayon de la boule pour la recherche des minima locaux est  $\rho^{(m)} = 2.5$  [[TH98], §6.3, p. 82].

## 3.2.8 Détecteur de Lowe : SIFT

### 3.2.8.1 La théorie

#### A) Le principe informel

David Lowe [Low04] propose de détecter des points qui sont extremums dans l'espace des échelles, donc invariants par changement d'échelle et rotation.

#### B) Les équations

Soient

- l'écart type  $\sigma \in \mathbb{R}^{+*}$ , encore dénommé échelle ;
- la fonction gaussienne  $G_{\sigma^2 Id_2}$  (équation (3.45)) ;
- le Laplacien  $\Delta$  d'une fonction :

$$\left\{ \begin{array}{lll} \Delta : \mathcal{C}^2(\mathbb{R}^2) & \rightarrow & \mathcal{C}^2(\mathbb{R}^2) \\ f & \mapsto & \Delta f : \mathbb{R}^2 \rightarrow \mathbb{R} \\ & & X \mapsto \Delta f(X) = \frac{\partial^2 f}{\partial x^2}(X) + \frac{\partial^2 f}{\partial y^2}(X) \end{array} \right. \quad (3.94)$$

Considérons alors :

- l'application  $h$  :

$$\left\{ \begin{array}{lll} h : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ X^\sigma = [X^T \sigma]^T & \mapsto & h(X^\sigma) = (\sigma^2 \Delta G_{\sigma^2 Id_2} \otimes I)(X) \\ & & = \left( \sigma^2 \left( \frac{\partial^2 G_{\sigma^2 Id_2}}{\partial x^2} + \frac{\partial^2 G_{\sigma^2 Id_2}}{\partial y^2} \right) \otimes I \right)(X) \end{array} \right. \quad (3.95)$$

- l'ensemble  $\mathcal{S}_1$  des points qui sont minima locaux stricts de  $h$  :

$$\mathcal{S}_1 = \{ \bar{X}^\sigma \in \mathbb{R}^3 \mid \exists B^1 \stackrel{n}{=} B_2(\bar{X}^\sigma, \bar{\rho}^1) \text{ telle que } \forall X^\sigma (\neq \bar{X}^\sigma) \in B^1, h(\bar{X}^\sigma) < h(X^\sigma) \} \quad (3.96)$$

- l'ensemble  $\mathcal{S}_2$  des pixels qui sont maxima locaux stricts de  $h$  :

$$\mathcal{S}_2 = \{ \bar{X}^\sigma \in \mathbb{R}^3 \mid \exists B^2 \stackrel{n}{=} B_2(\bar{X}^\sigma, \bar{\rho}^2) \text{ telle que } \forall X^\sigma (\neq \bar{X}^\sigma) \in B^2, h(X^\sigma) < h(\bar{X}^\sigma) \} \quad (3.97)$$

$\bar{X}^\sigma \in \mathcal{S}_1 \cup \mathcal{S}_2 \Rightarrow \nabla h(\bar{X}^\sigma) = 0$  (Équation d'Euler [[Cia85], théorème 7.2.1, p. 146]).

- l'ensemble  $\mathcal{S}_3$  des pixels qui sont extremums locaux stricts de  $h$  et qui possèdent un contraste suffisant :

$$\mathcal{S}_3 = \{ \bar{X}^\sigma \in \mathcal{S}_1 \cup \mathcal{S}_2 \text{ tel que } |h(\bar{X}^\sigma)| \geq \tau_{c(\text{contraste})} \} \quad (3.98)$$

- l'ensemble  $\mathcal{S}_4$  des pixels qui sont extremums locaux stricts de  $h$ , qui possèdent un contraste suffisant et qui se situent sur une frontière :

$$\mathcal{S}_4 = \{\bar{X}^\sigma \in \mathcal{S}_3 \text{ tel que } \frac{(Tr(\nabla_{\bar{X}}^2 h(\bar{X}^\sigma)))^2}{Det(\nabla_{\bar{X}}^2 h(\bar{X}^\sigma))} \leq \tau_{f(frontière)} = \frac{(r+1)^2}{r}\} \quad (3.99)$$

La justification de  $\mathcal{S}_4$  est de la même nature que pour Harris, c'est-à-dire aboutir à des valeurs propres de la matrice hessienne  $\nabla_{\bar{X}}^2 h(\bar{X}^\sigma)$  très proches en évitant leur calcul et éliminer le point lorsque  $r = \frac{\lambda_{max}}{\lambda_{min}} > 10$ .

L'ensemble  $\mathcal{S}^{\mathcal{L}}$  des points clefs selon Lowe [Low04] est  $\mathcal{S}^{\mathcal{L}} = \mathcal{S}_4$ .

$$\left\{ \begin{array}{ll} \mathcal{L} : \mathbb{N}_n \times \mathbb{N}_m \times \Sigma & \rightarrow \mathbb{R} \\ \bar{X}^\sigma & \mapsto \mathcal{L}(\bar{X}^\sigma) = \frac{(Tr(\nabla_{\bar{X}}^2 h(\bar{X}^\sigma)))^2}{Det(\nabla_{\bar{X}}^2 h(\bar{X}^\sigma))} \end{array} \right. \quad (3.100)$$

### 3.2.8.2 La pratique

La difficulté majeure consiste à passer de la formulation dans le domaine continu à l'implémentation dans le domaine discret et en particulier à gérer l'échantillonnage de l'espace des échelles  $\sigma$ .

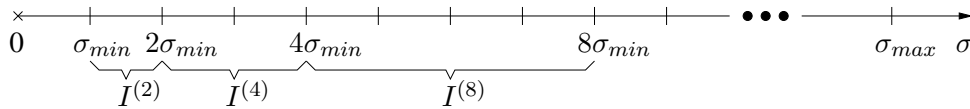
Dans ce cadre-là, Lowe propose les adaptations qui suivent.

#### A) Remplacement du Laplacien de Gaussienne normalisé $\sigma^2 \Delta G_{\sigma^2 Id_2}$ par une Différence de Gaussiennes (DOG)

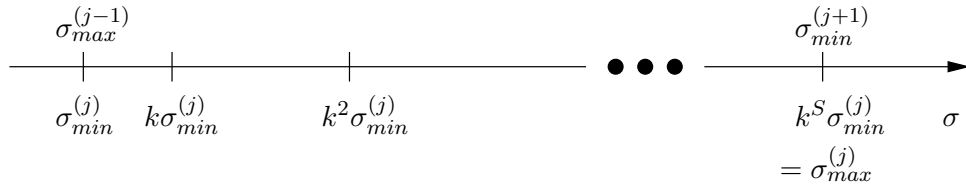
L'expression  $\sigma^2 \Delta G_{\sigma^2 Id_2}(X)$  devient  $DOG_{\sigma^2 Id_2}(X) = G_{(k\sigma)^2 Id_2}(X) - G_{\sigma^2 Id_2}(X)$ , où  $k$  est une constante multiplicative ne dépendant que de l'échantillonnage de l'espace des échelles.

#### B) Échantillonnage de l'espace des échelles « $\sigma$ »

- i) Comme illustré ci-après, l'intervalle  $[\sigma_{min}, \sigma_{max}]$  est divisé en octaves :



- 2i) Chaque intervalle  $I^{(2^j)} = [2^{j-1}\sigma_{min}, 2^j\sigma_{min}] \stackrel{n}{=} [\sigma_{min}^{(j)}, \sigma_{max}^{(j)}]$ ,  $j = 1, 2, \dots, O$  est échantillonné en  $(S+1)$  points aux positions :



de sorte que le rapport  $k$  qui lie deux  $\sigma$  successifs satisfait à  $k^S = 2 \Leftrightarrow k = 2^{\frac{1}{S}}$  (typiquement  $S = 3$ ). On obtient ainsi un sous-ensemble

$$\hat{\Sigma}(\subset \mathbb{R}) = \{\sigma_{min}, k\sigma_{min}^{(1)}, \dots, \sigma_{max}\}.$$

**C) Échantillonnage de l'espace  $(x, y)$** 

À partir d'une image brute  $m$  (lignes)  $\times n$  (colonnes), l'échantillonnage s'effectue ainsi :

- i) sur-échantillonnage éventuel de l'image brute pour passer à une image  $2m$  (lignes)  $\times 2n$  (colonnes) – via un polynôme interpolateur à deux indéterminées de degré global égal à 2 et de degrés partiels égaux à 1 – ;
- 2i) **Pour  $o$ (octave) de 1 à  $O_{max}$  Faire**  
échantillonner pour ne retenir qu'une image  $\lfloor \frac{M}{2^o} \rfloor$  (lignes)  $\times \lfloor \frac{N}{2^o} \rfloor$  (colonnes)  
formée des pixels aux positions  $[2^o, 2^{o+1}, \dots, \lfloor \frac{M}{2^o} \rfloor] \times [2^o, 2^{o+1}, \dots, \lfloor \frac{N}{2^o} \rfloor]$

**Fin Pour.**

**D) Détection des extremums**

Elle se fait en deux temps :

- i) La localisation au pixel près.

Les ensembles  $\mathcal{S}_1$  et  $\mathcal{S}_2$  sont remplacés respectivement par  $\hat{\mathcal{S}}_1$  et  $\hat{\mathcal{S}}_2$  avec :

$$\hat{\mathcal{S}}_1 = \{\hat{X}^\sigma \in \mathbb{Z}^2 \times \hat{\Sigma} \mid \forall X^\sigma (\neq \hat{X}^\sigma) \in \hat{B}_\infty^{(d)}(\hat{X}^\sigma, 1), h(\hat{X}^\sigma) < h(X^\sigma)\} \quad (3.101)$$

$$\hat{\mathcal{S}}_2 = \{\hat{X}^\sigma \in \mathbb{Z}^2 \times \hat{\Sigma} \mid \forall X^\sigma (\neq \hat{X}^\sigma) \in \hat{B}_\infty^{(d)}(\hat{X}^\sigma, 1), h(X^\sigma) < h(\hat{X}^\sigma)\} \quad (3.102)$$

où  $\hat{B}_\infty^{(d)}(\hat{X}^\sigma, 1) = B_\infty^{(f)}(\hat{X}^\sigma, 1) \cap (\mathbb{Z}^2 \times \hat{\Sigma})$ .

- 2i) La localisation à la fraction de pixel près [BL02].

Pour tout  $\hat{X}^\sigma$  dans  $(\hat{\mathcal{S}}_1 \cup \hat{\mathcal{S}}_2)$ , on exécute une itération de Newton sur le problème

$$\begin{cases} \min & h(X^\sigma) \\ X^\sigma \in \mathbb{R}^3 \end{cases}, \text{ où le point initial de l'algorithme est } (X^\sigma)_0 = \hat{X}^\sigma.$$

On aboutit à  $\hat{\hat{X}}^\sigma = \hat{X}^\sigma - (\nabla^2 h(\hat{X}^\sigma))^{-1} \nabla h(\hat{X}^\sigma)$ , d'où les ensembles  $\hat{\hat{\mathcal{S}}}_1$  et  $\hat{\hat{\mathcal{S}}}_2$ .

Ceci équivaut en termes d'interpolation à rechercher les extremums locaux du polynôme interpolateur de Hermite à deux indéterminées, de degré global égal à deux et de degrés partiels égaux à deux, connaissant :

$$\begin{cases} h(\hat{X}^\sigma) \\ \nabla h(\hat{X}^\sigma) \text{ soit } \begin{bmatrix} \frac{\partial h}{\partial x}(\hat{X}^\sigma) & \frac{\partial h}{\partial y}(\hat{X}^\sigma) & \frac{\partial h}{\partial \sigma}(\hat{X}^\sigma) \end{bmatrix}^T \\ \nabla^2 h(\hat{X}^\sigma) \text{ soit } \begin{bmatrix} \frac{\partial^2 h}{\partial x^2}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial x \partial y}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial x \partial \sigma}(\hat{X}^\sigma) \\ \frac{\partial^2 h}{\partial y \partial x}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial y^2}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial y \partial \sigma}(\hat{X}^\sigma) \\ \frac{\partial^2 h}{\partial \sigma \partial x}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial \sigma \partial y}(\hat{X}^\sigma) & \frac{\partial^2 h}{\partial \sigma^2}(\hat{X}^\sigma) \end{bmatrix} \end{cases}$$

**Remarques**

- i)  $h(\hat{\hat{X}}^\sigma) = h(\hat{X}^\sigma) + \frac{1}{2} (\nabla h(\hat{X}^\sigma))^T \hat{\hat{X}}^\sigma$
- 2i) Tous les calculs de dérivées (partielles) se font par différences finies.

### E) Élimination des extremums dans les zones à contraste faible

L'ensemble  $\mathcal{S}_3$  est remplacé très logiquement par  $\hat{\hat{\mathcal{S}}}_3$ .

### F) Élimination des extremums proches des contours

L'ensemble  $\mathcal{S}_4$  est remplacé très logiquement par  $\hat{\hat{\mathcal{S}}}_4$ .

#### 3.2.8.3 Remarque

Nous n'avons présenté que le détecteur de « points-clefs », pour reprendre la terminologie de son auteur. L'implémentation de D. Lowe [Low04] fournit pour chaque « points-clefs » un descripteur de 128 caractéristiques, approprié à faire de la mise en correspondance entre deux images d'une même scène dès lors que l'on définit une distance [[NSB07], §3.1, p. 514].

## 3.2.9 Détecteur de Mikolajczyk et Schmid

### 3.2.9.1 La théorie

#### A) Le principe informel

L'objectif est de rechercher les points d'intérêt au sens de  $\mathcal{HS}^{smb}$  (équation (3.62)) (version améliorée de Schmid *et al.* donc paramétrés par  $\bar{\sigma}_D$ , l'écart type dit de différentiation, et  $\bar{\sigma}_I$ , l'écart type dit d'intégration). Ces points d'intérêt sont pressentis être invariants aux changements d'échelle.

Dans sa version simplifiée (« Harris-Laplace » [[MS04], p. 68]), un point d'intérêt  $(\bar{X}, \bar{\sigma}_I)$  au sens de Mikolajczyk et Schmid ( $\mathcal{MS}$ ) satisfait à :

- a)  $\bar{X}$  est un maximum local du critère  $\mathcal{HS}^{smb}$  paramétré par l'échelle  $\bar{\sigma}_I$
- b)  $\bar{\sigma}_I$  est l'écart type qui maximise localement la fonction  $h$  (équation (3.95)) pour  $X = \bar{X}$  fixé, conformément aux conclusions de Bretzner et Lindeberg [[BL98], éq. (8), p. 5] et de Lindeberg [Lin98] concernant les détecteurs de « pâtés » (*blobs*, en anglais).

Notons  $\mathcal{S}^{\mathcal{MS}}$  l'ensemble des points d'intérêt de Mikolajczyk et Schmid.

#### B) Les équations

##### B.1) Détermination des points d'intérêt de Harris-Stephens à l'échelle $\sigma_I$

Cette détermination est présentée dans l'algorithme 3.2.2.

**Pour tout**  $\sigma_I \in \Sigma_I = \{\sigma_I^{min}, \dots, \sigma_I^{max}\}$  **Faire**  
**Pour tout**  $X \in \mathbb{N}_n \times \mathbb{N}_m$  **Faire**  
 Calculer  $\mathcal{HS}^{smb}(X^{\sigma_I})$   
**Fin Pour**  
 Déterminer les points  $\bar{X}$  correspondant aux maxima locaux du critère  $\mathcal{HS}^{smb}$  :  
 $\{\bar{X}^{\sigma_I}\}$   
**Fin Pour**

Algorithme 3.2.2: Algorithme de détermination des points d'intérêt de Harris-Stephens à l'échelle  $\sigma_I$

L'ensemble  $\Sigma_I$  (respectivement  $\Sigma_D$ ) est constitué des valeurs :

$$\sigma_I^n = \xi^n \sigma_I^0, n = 0, \dots, \text{ où } \xi = \sqrt{2} \approx 1.4$$

$$(\text{respectivement } \sigma_D^n = s \sigma_I^n, \text{ où } s = \frac{\xi}{2} = \frac{\sqrt{2}}{2} \approx 0.7).$$

### B.2) Détermination des $\bar{\sigma}_I$ et simultanément des points d'intérêt de précision optimale $\bar{\bar{X}}^{\sigma_I}$

Cette détermination est présentée dans l'algorithme 3.2.3.

**Pour tout**  $\bar{X}^{\sigma_I}$  **Faire**  
 aberrant  $\leftarrow$  FAUX  
 $k \leftarrow 0$   
 $X^k \leftarrow \bar{X}^{\sigma_I}$   
 $\sigma_I^k \leftarrow \sigma_I$   
**Répéter**  
 $\sigma_I^{k+1} \leftarrow \begin{cases} \text{argmax} & h_{|\{X^k\} \times \Sigma_I^{k+1}|}([X^k]^T \sigma_I^k)^T \text{ (cf éq. (3.95))} \\ \sigma_I \in \Sigma_I^{k+1} \end{cases} \quad (\text{A-3.2.3-1})$   
**Si**  $(\sigma_I^{k+1} \neq \sigma_I^k)$  **Alors**  
 $X^{k+1} \leftarrow \begin{cases} \text{argmax} & \mathcal{HS}^{smb}([X^T \sigma_I^{k+1}]^T) \\ X \in B_2(X^k, \rho^{(ms)}) \end{cases} \quad (\text{A-3.2.3-2})$   
**Sinon Si**  $(\sigma_I^{k+1} = 0)$  **Alors**  
 aberrant  $\leftarrow$  VRAI  
**Fin Si**  
 $k \leftarrow k + 1$   
**Jusqu'à** ((aberrant) ou  $(X^k = X^{k-1}$  et aussi  $\sigma_I^k = \sigma_I^{k-1}$ ))  
**Si** (non(aberrant)) **Alors**  
 $\bar{\bar{X}}^{\sigma_I} \leftarrow (X^k, \sigma_I^k)$   
**Fin Si**  
**Fin Pour**

Algorithme 3.2.3: Algorithme de détermination simultanée des  $\bar{\sigma}_I$  et des points d'intérêt  $\bar{\bar{X}}^{\sigma_I}$  de précision optimale

### B.3) La version simplifiée

Elle consiste simplement à rejeter des points d'intérêt  $\bar{X}^{\sigma_I} = [\bar{X}^T \sigma_I]^T$  issus de l'étape (B.1) qui ne satisfont pas à :

$$\begin{cases} h(\bar{X}^{\sigma_I}) > h([\bar{X}^T \sqrt{2} \sigma_I]^T) \\ h(\bar{X}^{\sigma_I}) > h([\bar{X}^T \frac{\sigma_I}{\sqrt{2}}]^T) \end{cases} . \quad (3.103)$$

#### 3.2.9.2 La pratique

L'ensemble  $\Sigma_I^{k+1}$  est choisi tel que :

$$\Sigma_I^{k+1} = \{\xi' \sigma_I^k\}_{\xi' \in \{\frac{\sqrt{2}}{2}(\xi'')^n\}_{n \in [0,7]_{\mathbb{N}}}} \quad \text{avec } \xi'' = \sqrt[7]{2} \approx 1.104,$$

d'où :

$$\Sigma_I^{k+1} \approx \{t \sigma_I^k\}_{t \in \{0.7, 0.8, \dots, 1.1, \dots, 1.4\}} \quad (3.104)$$

#### 3.2.9.3 Comparaison avec SIFT

- i) Les deux opérateurs de Lowe [Low04] et de Mikolajczyk et Schmid [MS04] s'appuient sur l'évaluation de la fonction  $h$  (issue des travaux de Lindeberg) porteuse d'informations multi-échelles intrinsèques à toute image numérique.
- 2i) Il n'existe aucune relation d'inclusion entre  $\mathcal{S}^{\mathcal{L}}$ , l'ensemble des points clefs de Lowe, et  $\mathcal{S}^{\mathcal{MS}}$ , l'ensemble des points d'intérêt de Mikolajczyk et Schmid, car même en confondant Laplacien de Gaussienne et Différence de Gaussiennes :
  - a) dans le premier cas, un point  $(\bar{X}^{\sigma})^{\mathcal{L}}$  est extremum local de la fonction  $h$  des trois variables  $(x, y, \sigma)$ ,
  - b) dans le second cas, un point  $(\bar{X}^{\bar{\sigma}_I})^{\mathcal{MS}}$  satisfait à :

$$\begin{cases} \bar{X} \text{ est un point de } \mathcal{HS}^{smb} \\ \bar{\sigma}_I \text{ rend seulement } h_{|\{\bar{X}\} \times \ll \mathbb{R} \gg} \text{ maximale} \end{cases} .$$

- 3i) L'opérateur SIFT ignore tout des caractéristiques des points d'intérêt de Harris.
- 4i) L'opérateur  $\mathcal{MS}$  ignore tout des propriétés de la matrice hessienne  $\nabla_X^2 h(\bar{X}^{\sigma})$  (équation (3.99)), donc de la « qualité » d'un extremum de  $h_{|\mathbb{R}^2 \times \{\bar{\sigma}_I\}}$ .

### 3.2.10 Détecteur de Mikolajczyk et Schmid invariant aux transformations affines

Il s'agit d'une extension des points d'intérêt invariants aux changements d'échelle aux points d'intérêt invariants aux transformations affines (plus précisément linéaires).

#### 3.2.10.1 La théorie

##### A) Le principe informel

L'idée directrice est :



- i) de remplacer les noyaux de convolution 2D, à symétrie circulaire (c'est-à-dire dont le support est un disque), par des noyaux à support elliptique ;
- 2i) d'appliquer la même démarche que pour les points d'intérêt invariants aux changements d'échelle.

## B) Les équations

### B.1) Les gaussiennes adaptées au traitement des transformations affines

Les applications gaussiennes d'intérêt sont  $G_{C_I}$  et  $G_{C_D}$  (voir équation (3.45)), où les matrices de variances-covariances d'intégration et de différentiation sont respectivement  $C_I = \sigma_I^2 Id_2$  et  $C_D = \sigma_D^2 Id_2$ .

### B.2) Les difficultés

Même en liant la matrice de variances-covariances  $C^I$  à  $C^D$  (comme pour le cas  $\sigma_I$  et  $\sigma_D$ ), il est difficile d'échantillonner l'espace produit  $\Sigma_I \times \Sigma_D$ , où :

- $\Sigma_I = \{C_I \in M_2(\mathbb{R}) | C_I = C_I^T > 0\}$  (donc  $C_I$  est symétrique définie positive, d'où  $C_I \in S_2(\mathbb{R})$ ),
- $\Sigma_D = \{C_D \in S_2(\mathbb{R})\}$ .

### B.3) La solution de Mikolajczyk et Schmid

Ils estiment à l'aide d'une procédure itérative la transformation affine (seul le terme linéaire est traité) et son inverse de manière à se ramener au cas des points invariants par changement d'échelle et rotation.

- i) La détermination des points de  $\mathcal{HS}^{smb}$  à l'échelle  $\sigma_I$ .  
Les points  $\bar{X}^{\sigma_I}$  sont issus de l'algorithme 3.2.2.
- 2i) La détermination des  $\bar{C}^I$  et simultanément des points d'intérêt  $\bar{\bar{X}}$  de précision optimale (voir algorithme 3.2.4).

**Pour tout  $\bar{X}^{\sigma_I}$  Faire**

$k \leftarrow 0$

$T^{(0)} \leftarrow Id_2$

$X^{(0)} \leftarrow \bar{X}^{\sigma_I}$

**Répéter**

$I_T(X_T) \leftarrow I(T^{(k)} X_T), \forall X_T \in B_2(\underbrace{(T^{(k)})^{-1} X^{(k)}}_{\hat{X}_T}, \rho^{(k)})$

Déterminer  $\bar{\sigma}_I^{(T)}$  au point  $\hat{X}_T$  - cf. algorithme 3.2.3, action (A-3.2.3-1)

Déterminer  $\bar{\sigma}_D^{(T)} = \begin{cases} \text{argmax} & \frac{\lambda_{\min}(\hat{M}^{hs})}{\lambda_{\max}(\bar{M}^{hs})} \\ \sigma_D = s\bar{\sigma}_I^{(T)} | s \in \{0.5, 0.55, \dots, 0.75\} \end{cases}$ ,

où  $\hat{M}^{hs} = M^{hs+smb}(\hat{X}_T, \bar{\sigma}_I^{(T)}, \sigma_D)$

Déterminer  $\bar{X}_T$  - cf. algorithme 3.2.3, action (A-3.2.3-2)

$X^{(k+1)} \leftarrow T^{(k)} \bar{X}_T$  - il y a peut-être une translation à faire

$\bar{M}^{hs} \leftarrow (M^{hs+smb}(\bar{X}_T, \bar{\sigma}_I^{(T)}, \bar{\sigma}_D^{(T)}))^{-\frac{1}{2}}$  (A-3.2.4-1)

$T^{(k+1)} \leftarrow \bar{M}^{hs} T^{(k)}$  (A-3.2.4-2)

$T^{(k+1)} \leftarrow \frac{T^{(k+1)}}{\lambda_{\max}(T^{(k+1)})}$  (A-3.2.4-3)

Déterminer  $\rho^{(k+1)}$

$k \leftarrow k + 1$

**Jusqu'à**  $\left(\frac{\lambda_{\min}(\bar{M}^{hs})}{\lambda_{\max}(\bar{M}^{hs})} > -\tau_{\bar{M}^{hs}}^{(min)} + 1\right)$  ou  $\left(\frac{\lambda_{\max}(\bar{M}^{hs})}{\lambda_{\min}(\bar{M}^{hs})} > \tau_{\bar{M}^{hs}}^{(max)}\right)$

**Si**  $\left(\frac{\lambda_{\max}(\bar{M}^{hs})}{\lambda_{\min}(\bar{M}^{hs})} \leq \tau_{\bar{M}^{hs}}^{(max)}\right)$  **Alors**

$\bar{X}^{\sigma_I} \leftarrow X^{(k)}$

**Fin Si**

**Fin Pour**

Algorithme 3.2.4: Détermination simultanée des  $\bar{C}^I$  et des points d'intérêt  $\bar{X}$  de précision optimale

La justification des actions (A-3.2.4-1), (A-3.2.4-2) et (A-3.2.4-3) de l'algorithme 3.2.4 est donnée par Lindeberg et Gårding [LG01] (déjà implémentée par Baumberg [[Bau00], §4.2, p. 4]). En voici les étapes essentielles.

Soient :

- a)  $I^1$  et  $I^2$  deux images liées géométriquement par une application linéaire inversible  $T$  (l'abstraction est faite de la translation de la transformation affine), à savoir :

$$\forall X \in \mathbb{N}_n \times \mathbb{N}_m, I^1(X) = I^2(TX). \quad (3.105)$$

- b)  $\tilde{I}_{C_1}^1$  et  $\tilde{I}_{C_2}^2$  les représentations affines « espace-échelle » au sens de Lindeberg et Gårding [[LG01], éq. (27) et (28)], où  $C_1, C_2 \in S_2(\mathbb{R})$  :

$$\tilde{I}_{C_1}^1 = G_{C_1} \otimes I^1 \quad \tilde{I}_{C_2}^2 = G_{C_2} \otimes I^2 \quad (3.106)$$

Alors :

- a)  $\forall C_1 \in S_2(\mathbb{R}), \tilde{I}_{C_1}^1 = \tilde{I}_{C_2}^2 \circ T$  [[LG01], éq. (29)], où  $C_2 = TC_1 T^T$  [[LG01], éq. (30)],

- b) En appelant  $M^{hs+smb}(X, C_I, C_D)$  la matrice  $2 \times 2$ , chère à Harris et Stephens et Schmid *et al.* (équation (3.64)) :

$$M^{(1)}(X, C_I, C_D) = T^T (M^{(2)}(TX, TC_I T^T, TC_D T^T)) T \quad (3.107)$$

[[LG01], éq. (45)] et [[MS04], éq. (5), p. 70], où  $M^{(1)}$  (respectivement  $M^{(2)}$ ) est calculée sur  $I^1$  (respectivement  $I^2$ ). L'équation (45) de Lindeberg [LG01] est celle qui permet d'estimer  $T$  connaissant  $M^{(1)}$  et  $M^{(2)}$  liées par :

$$M^{(1)} = T^T M^{(2)} T \quad [[LG01], \text{éq. (17)}]. \quad (3.108)$$

En effet :

$$T = (M^{(2)})^{-\frac{1}{2}} O (M^{(1)})^{\frac{1}{2}} \quad [[LG01], \text{éq. (14)}], \quad (3.109)$$

où  $OO^T = O^T O = Id_2$  est l'unique solution de l'équation (3.108).

Voici notre preuve personnelle de ce résultat (qui a été approuvée par Tony Lindeberg, surpris de ne pas la retrouver explicitement dans ses publications) :

- i) Rappelons que,  $\forall A \in S_n(\mathbb{R})$ ,  $\exists O_A (O_A^T O_A = O_A O_A^T = Id_n)$  et  $\Lambda_A = \text{diag}(\lambda_i)_{i=1,\dots,n}$  satisfaisant à :

$$A = O_A^T \Lambda_A O_A = (O_A^T \Lambda_A^{\frac{1}{2}} O_A) (O_A^T \Lambda_A^{\frac{1}{2}} O_A),$$

où  $\Lambda_A^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_i})_{i=1,\dots,n}$ .

- 2i)  $T$  donnée par l'équation (3.109) vérifie bien l'équation (3.108) :

$$M^{(1)} = (M^{(1)})^{\frac{1}{2}} O^T (M^{(2)})^{-\frac{1}{2}} M^{(2)} (M^{(2)})^{-\frac{1}{2}} O (M^{(1)})^{\frac{1}{2}}$$

- 3i)  $O$  est unique et a pour expression :

$$O = (M^{(2)})^{\frac{1}{2}} T (M^{(1)})^{-\frac{1}{2}}$$

- 4i)  $O$  est bien orthogonale :

$$O^T O = O O^T = (M^{(1)})^{-\frac{1}{2}} T^T (M^{(2)})^{\frac{1}{2}} (M^{(2)})^{\frac{1}{2}} T (M^{(1)})^{-\frac{1}{2}} = Id_2$$

- 5i)  $\text{Det}(O) = \frac{\sqrt{\text{Det}(M^{(2)})}}{\sqrt{\text{Det}(M^{(1)})}} \text{Det}(T) = \pm 1$  car, d'après l'équation (3.108),

$$\text{Det}(M^{(1)}) = (\text{Det}(T))^2 \text{Det}(M^{(2)})$$

- 6i) On est en dimension 2, donc  $O$  est soit une matrice de rotation  $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ , soit une symétrie par rapport à l'axe  $y'Oy$   $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ , soit une symétrie par rapport à l'axe  $x'Ox$   $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$  [[WD04], théorème 32, p. 611].

c) Si

$$\begin{cases} C_I = \sigma_I (M^{(1)}(X, C_I, C_D))^{-1} \\ C_D = \sigma_D (M^{(1)}(X, C_I, C_D))^{-1} \end{cases}, \quad (3.110)$$

alors :

$$\begin{cases} TC_I T^T = \sigma_I (M^{(2)}(TX, TC_I T^T, TC_D T^T))^{-1} \\ TC_D T^T = \sigma_D (M^{(2)}(TX, TC_I T^T, TC_D T^T))^{-1} \end{cases} \quad (3.111)$$

[[LG01], éq. (47), (48), (49), (50)], [[MTS<sup>+</sup>05], éq. (7) et (8)]. Ceci signifie que si des points d'intérêt satisfont aux conditions suffisantes (équation (3.110)), alors ces propriétés sont conservées lors d'une transformation linéaire. Ces points d'intérêt sont considérés invariants aux transformations linéaires.

d) En appelant :

$$\begin{cases} \tilde{X} &= (M^{(1)})^{\frac{1}{2}} X \\ \text{et} & \\ \tilde{X}^T &= (M^{(2)})^{\frac{1}{2}} TX \end{cases},$$

alors l'équation (3.109) implique  $\tilde{X}^T = O\tilde{X}$  ([MS04], éq. (9)) et [Bau00], fig. 2, p. 5] ce qui montre que la transformation linéaire  $T$  est ramenée à une rotation ou une symétrie par rapport à  $(x'Ox)$  (ou  $(y'Oy)$ ).

### 3.2.10.2 La pratique

Il n'y a rien à rajouter par rapport à ce qui a déjà été dit.

## 3.3 Détecter pour suivre

### 3.3.1 Shi et Tomasi : KLT

Comme cela a déjà été indiqué dans la section 3.2.5, les points d'intérêt de Shi et Tomasi [ST94] se veulent optimaux pour leur algorithme de suivi nommé KLT (Kanade-Lucas-Tomasi). La vocation de KLT est de suivre  $n_i \geq 1$  points d'intérêt indépendamment les uns des autres.

### 3.3.2 Zivkovic et Van der Heijden

#### 3.3.2.1 Principe informel

L'algorithme de Zivkovic et Van der Heijden (ZV) [ZdH02, ZdH04] aboutit à un classement d'un ensemble de points d'intérêt préalablement détectés. Zivkovic et Van der Heijden tirent eux aussi profit de la célèbre matrice de Harris et Stephens  $M^{hs}$  (équation (3.48)) intervenant dans (KLT) et définissent pour **région de convergence** d'un point  $\bar{X}$ ,  $\mathcal{R}(\bar{X}) = B_2(O, \rho^{\bar{X}})$  constituée de « tous » les déplacements traités avec succès par la première itération de KLT. Toute la difficulté réside dans le calcul de  $\rho^{\bar{X}}$  : la réponse de Zivkovic et Van der Heijden est une heuristique consistant à ne tester qu'un sous-ensemble des déplacements de  $\mathcal{R}(\bar{X})$  [[ZdH04], §4, p. 6]. Qualitativement, la région de convergence est petite dans les zones homogènes et grande dans les zones fortement texturées.

## 3.3.2.2 Algorithme (ZV)

échecKLT  $\leftarrow$  FAUX  
 $\Delta\beta^{(a)} \leftarrow \Delta\beta_0^{(a)}$  - - perturbation artificielle initiale  
**Répéter**  
   Construire l'image artificielle  $I^{(a)}$  :  
     
$$\forall \mathbf{m} \in \mathbb{N}_n \times \mathbb{N}_m, I^{(a)}(\mathbf{m}) \leftarrow I_{ref}(T^{-1}(\mathbf{m}, \Delta\beta^{(a)}))$$
  
   Simuler la **première itération** de l'algorithme KLT sur  $I^{(a)}$   
   **Si** elle converge **Alors**  
     Augmenter l'amplitude de la perturbation  $\Delta\beta^{(a)}$   
   **Sinon**  
     échecKLT  $\leftarrow$  VRAI  
   **Fin si**  
**Jusqu'à échecKLT**

Algorithme 3.3.1: Le principe de Zivkovic et Van der Heijden.

L'amplitude  $\rho^{\bar{X}}$  de la région de convergence au sens de Zivkovic et Van der Heijden est définie par la norme de la plus grande perturbation artificielle  $\Delta\beta^{(a)}$  gagnante. Les points sélectionnés sont des maxima locaux de :

$$ZV(\bar{X}) = \rho^{\bar{X}}. \quad (3.112)$$

## 3.4 Sélection de sous-ensembles

Voici deux méthodes :

- la sélection séquentielle par ajout, fondée sur les statistiques ;
- la sélection d'un sous-ensemble (en anglais, *subset selection*), fondée sur l'algèbre linéaire.

## 3.4.1 Sélection séquentielle par ajout : approche statistique

## 3.4.1.1 Matrice de variances-covariances des paramètres d'un problème aux moindres carrés linéaires

Considérons le modèle linéaire statistique générique :

$$A\beta = b + \epsilon, \quad (3.113)$$

où :

- $A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \in \mathbb{R}^{m \times n}$ , de rang  $n$ , est composée des  $m$  variables explicatives  $a_i$  ;
- $\beta \in \mathbb{R}^n$  est le vecteur des paramètres inconnus ;
- $b \in \mathbb{R}^m$  est composé des  $m$  réponses ;
- $\epsilon \in \mathbb{R}^m$  est un bruit blanc gaussien ( $\epsilon \sim \mathcal{N}(0, \sigma^2 Id_m)$ ).

Alors, par le théorème de Gauss-Markoff [[Bjö96], théorème 1.1.1, p. 4], [[Lue69], théorème 1, p. 86], la solution  $\hat{\beta} = (A^T A)^{-1} A^T b$  du problème aux moindres carrés linéaires :

$$\begin{cases} \min & \frac{1}{2} \|A\beta - b\|_2^2 \\ \beta \in \mathbb{R}^n \end{cases}$$

est un estimateur linéaire de  $\beta$ , non biaisé et de variance minimale :

$$V(\in \mathbb{R}^{n \times n}) = \sigma^2 C, \quad (3.114)$$

où :

$$C = (A^T A)^{-1}. \quad (3.115)$$

Un estimateur non biaisé de  $\sigma^2$  est :

$$\widehat{\sigma^2} = \frac{1}{2} \frac{\|A\hat{\beta} - b\|_2^2}{m - n} \quad (3.116)$$

### 3.4.1.2 Une mesure de l'incertitude des paramètres estimés

Pour mesurer la qualité des paramètres estimés, Reeves et Hezar [RH95] ainsi que Dellaert et Collins [DC99] s'appuient sur  $Tr(V)$  [[DC99], éq. (9), p. 7], alors que Jaakkola [Jaa02] considère  $Det(V)$ .

L'apprentissage actif consiste à ne garder qu'un nombre donné  $k$  parmi les  $m$  variables explicatives qui minimisent l'incertitude des paramètres estimés, à savoir  $Det(V)$ .

Remarquons que, si  $Det(V) \neq 0$  :

$$\begin{cases} Det(V^{-1}) = \frac{1}{Det(V)} \\ V^{-1} = \frac{1}{\sigma^2} A^T A \end{cases} \Rightarrow \begin{cases} \min & Det(V) \\ V \in \mathbb{R}^{n \times n} \end{cases} = \begin{cases} \max & Det(V^{-1}) \\ V \in \mathbb{R}^{n \times n} \end{cases},$$

ce qui évite le calcul de l'inverse de  $A^T A$ .

### 3.4.1.3 L'heuristique SFS

Le calcul exhaustif des valeurs  $Det(V^{-1})$ , associées aux  $\binom{m}{k}$  sous-ensembles de variables explicatives, est impossible dans la pratique :  $\binom{1000}{501} > 10^{299} !!!$

La sélection séquentielle par ajout (*Sequential Forward Selection* : SFS), due à Reeves et Hezar [RH95], est une heuristique permettant de construire, à partir d'un ensemble initial  $\Gamma^{(i)} = \{a_1, \dots, a_m\}$  de  $m$  variables explicatives  $a_i$ , un ensemble final  $\Gamma^{(f)} = \{a_{l_1}, \dots, a_{l_k}\}$  de  $k$  variables explicatives ( $n \leq k \leq m$ ), en  $(k - n)$  itérations, d'où le schéma itératif appliqué au problème :

$$\begin{cases} \max & Det(V^{-1}) \\ V \in \mathbb{R}^{n \times n} \\ V^{-1} = \frac{1}{\sigma^2} A^T A \\ A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \end{cases}.$$

**A) Concernant le schéma itératif**

$$\begin{cases} \Gamma_n^{(f)} = \{a_{l_1}, \dots, a_{l_n}\} \\ \Gamma_j^{(f)} = \Gamma_{j-1}^{(f)} \cup \{a_{l_j}\}, \forall j \in [n+1, k]_{\mathbb{N}} \end{cases}, \quad (3.117)$$

où :

$$a_{l_j} = \left\{ \operatorname{argmax}_{a \in \Gamma^{(i)} \setminus \Gamma_{j-1}^{(f)}} F_{SFS}^j(a) \right., \quad (3.118)$$

$$F_{SFS}^j(a) = \operatorname{Det} \left( \frac{1}{\sigma^2} \begin{bmatrix} A^{j-1} \\ a^T \end{bmatrix}^T \begin{bmatrix} A^{j-1} \\ a^T \end{bmatrix} \right), \quad (3.119)$$

$$\sigma^2 = \frac{1}{2} \frac{\left\| \begin{bmatrix} A^{j-1} \\ a^T \end{bmatrix} \hat{\beta} - \begin{bmatrix} b^{j-1} \\ \tilde{b} \end{bmatrix} \right\|_2^2}{j - n}, \quad (3.120)$$

$$\hat{\beta} = \begin{cases} \operatorname{argmin}_{\beta \in \mathbb{R}^n} \frac{1}{2} \left\| \begin{bmatrix} A^{j-1} \\ a^T \end{bmatrix} \beta - \begin{bmatrix} b^{j-1} \\ \tilde{b} \end{bmatrix} \right\|_2^2 \end{cases} \quad (3.121)$$

$$\begin{cases} A^n = \begin{bmatrix} a_{l_1}^T \\ \vdots \\ a_{l_n}^T \end{bmatrix} & \text{et } b^n = \begin{bmatrix} b_{l_1} \\ \vdots \\ b_{l_n} \end{bmatrix} \\ A^j = \begin{bmatrix} A^{j-1} \\ a_{l_j}^T \end{bmatrix} & \text{et } b^j = \begin{bmatrix} b^{j-1} \\ b_{l_j} \end{bmatrix}, \forall j \in [n+1, k]_{\mathbb{N}} \end{cases}. \quad (3.122)$$

**B) Concernant le choix de  $\Gamma_n^{(f)}$** 

Aucun des auteurs concernés ne propose de solution. Nous suggérons de suivre en l'adaptant, l'heuristique défendue et mise en œuvre par Rousseeuw et ses collaborateurs dans les logiciels :

- PROGRESS [[RH97], action 1, p. 7, action 2, p. 8] ;
- FAST\_LTS [[RD99a], §3.1, p. 6] ;
- FAST\_MCD [[RD99b], §4.1, p. 8].

Rappelons-en les grandes lignes.

```

Det_Max ← 0
Pour  $i$  de 1 à  $i_{max}$  Faire
  Choisir aléatoirement  $\Gamma_{temp} = \{a_{l_1}^i, \dots, a_{l_n}^i\}$ 
  Calculer  $\hat{\beta}^i$ , solution du système linéaire  $A^n \beta = b^n$ 
   $(V^n)^{-1} \leftarrow \frac{1}{\sigma^2} (A^n)^T A^n$ 
  Si  $\operatorname{Det}((V^n)^{-1}) > \operatorname{Det\_Max}$  Alors
     $\Gamma_n^{(f)} \leftarrow \Gamma_{temp}$ 
     $\operatorname{Det\_Max} \leftarrow \operatorname{Det}((V^n)^{-1})$ 
  Fin Si
Fin Pour

```

$i_{max}$  représente le nombre de tirages aléatoires de  $k$  variables explicatives parmi  $m$  (sous l'hypothèse :  $\frac{k}{m} \approx 0$ ) tel que  $P \geq \tau_P (\geq 0.95)$ , où :

- $P$  est la probabilité d'avoir un échantillon constitué de  $k$  variables explicatives non contaminées, à savoir :

$$P \approx 1 - (1 - (1 - \epsilon)^k)^{i_{max}},$$

- $\epsilon$  est le pourcentage de variables explicatives contaminées [[RL87], éq. (1.2), p. 198].

---

#### Remarque

La sélection séquentielle par retrait SBS (en anglais, *Sequential Backward Selection*) suit le même principe mais, dans ce cas, les variables explicatives sont retirées séquentiellement.

---

#### 3.4.1.4 Notre problème

En régression linéaire, il est des situations où on dispose de plusieurs modèles pour un même ensemble de variables explicatives. On passe d'un modèle plus « riche en paramètres » à un modèle moins « riche » (et vice-versa) par retrait (et vice-versa, par ajout) de colonnes dans la matrice  $A$  du problème aux moindres carrés linéaires [[Bjö96], §3.1.1, p. 127].

La mise à jour positive (« updating ») (respectivement négative (« downdating »)) est une méthode utilisée lors de l'ajout (respectivement du retrait) de lignes ou de colonnes, c'est-à-dire de variables explicatives ou de paramètres, dans le problème aux moindres carrés linéaires afin de :

- déterminer l'estimation  $\hat{\beta}$  du nouveau problème implicite,
- réduire les calculs à leur strict minimum, à savoir tirer profit des résultats hérités du problème d'origine.

Dans le cadre de la sélection de sous-ensembles de points d'intérêt, notre besoin concerne le choix de  $k$  paramètres associés au modèle de régression linéaire.

Vue sous l'angle de l'élimination des paramètres d'un modèle sur-paramétré (heuristique SBS), une démarche de bon sens conduit à « oublier », en priorité, les composantes  $\hat{\beta}_l$  de variance estimée maximale. Nous visons cet objectif en adaptant l'heuristique SFS (voir §3.4.1.3).

#### A) Concernant le schéma itératif

Reprenons la trame de la section 3.4.1.3 à partir de :

- $\Gamma^{(i)} = \{c_1, \dots, c_n\}$ , où  $c_j \in \mathbb{R}^m$ ,
- $A = \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix}$ ,
- $\Gamma^{(f)} = \{c_{l_1}, \dots, c_{l_k}\}$ ,  $k \leq n$ .

Alors :

$$\begin{cases} \Gamma_0^{(f)} = \emptyset \\ \Gamma_j^{(f)} = \Gamma_{j-1}^{(f)} \cup \{c_{l_j}\}, \forall j \in [1, k]_{\mathbb{N}} \end{cases}, \quad (3.123)$$

où :



$$c_{l_j} = \left\{ \operatorname{argmax}_{c \in \Gamma^{(i)} \setminus \Gamma_{j-1}^{(f)}} F_{SFS}^j(c) \right\}, \quad (3.124)$$

$$F_{SFS}^j(c) = \operatorname{Det} \left( \frac{1}{\sigma^2} \begin{bmatrix} A^{j-1} & c \end{bmatrix}^T \begin{bmatrix} A^{j-1} & c \end{bmatrix} \right), \quad (3.125)$$

$$\sigma^2 = \frac{1}{2} \frac{\| \begin{bmatrix} A^{j-1} & c \end{bmatrix} \hat{\beta} - b \|_2^2}{m - j}, \quad (3.126)$$

$$\hat{\beta} = \left\{ \operatorname{argmin}_{\beta \in \mathbb{R}^j} \frac{1}{2} \left\| \begin{bmatrix} A^{j-1} & c \end{bmatrix} \beta - b \right\|_2^2 \right\}, \quad (3.127)$$

$$\begin{cases} A^0 = [ ] \\ A^j = \begin{bmatrix} A^{j-1} & c_{l_j} \end{bmatrix} \end{cases} \quad (3.128)$$

## B) Concernant notre application

### B.1) Une complication de plus !

Les paramètres à sélectionner sont ici les composantes  $a_{ij}$  du prédicteur linéaire  $A$  de la célèbre relation «  $A\Delta\mathbf{I} = \Delta\beta$  » qui est au cœur de l'algorithme de suivi de Jurie et Dhome.

Rappelons que l'estimation de  $A$  (voir §2.5.1.2) passe par la résolution de  $p$  (nombre de paramètres du modèle de mouvement, par exemple 8 pour une homographie) problèmes aux moindres carrés linéaires partageant la même matrice «  $A$  », à savoir (voir l'équation (2.109)) :

$$\forall q = 1..p, \widehat{a}_q = \left( \left\{ \operatorname{argmin}_{a \in \mathbb{R}^m} \| C^T a - b_q^T \|_2^2 \right\} \right)^T \quad (3.129)$$

Cette contrainte induit une difficulté supplémentaire. En effet, l'application de l'heuristique SFS à chacun des  $p$  problèmes fournit a priori  $p$  paramètres différents. La question est donc : quel paramètre (donc quelle colonne de  $C^T$ ) commun aux  $p$  problèmes retenir à chaque itération de l'heuristique SFS ? Notre réponse est : celui qui minimise le critère LTS (moindres carrés tronqués à  $h = \frac{\alpha}{100}p$  ( $\alpha > 50$ )) [[RL87], éq. (2.16), p. 15] établi à partir des  $p$  incertitudes.

---

#### Remarque

Souvenons-nous que chaque colonne  $c_l$  de  $C^T$  est en « bijection » avec le point d'intérêt  $\mathbf{m}_l$  puisque :

$$c_l = \Delta\mathbf{I}(\mathbf{m}_l) (\in \mathbb{R}^{N_a}) = \begin{bmatrix} I_{ref}(T(\mathbf{m}_l, \beta_0 + \Delta\beta^1)) - I_{ref}(T(\mathbf{m}_l, \beta_0)) \\ \vdots \\ I_{ref}(T(\mathbf{m}_l, \beta_0 + \Delta\beta^{N_a})) - I_{ref}(T(\mathbf{m}_l, \beta_0)) \end{bmatrix}. \quad (3.130)$$


---

## B.2) Les équations

Reprenons la trame de la section 3.4.1.3 en partant de :

- $\Omega^{(i)} = \{\mathbf{m}_1, \dots, \mathbf{m}_n\}$ , composé de  $n$  points d'intérêt, auquel est associé  $\Gamma^{(i)} = \{c_1, \dots, c_n\}$ , où  $c_j \in \mathbb{R}^m$ ,
- $A = \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix}$ ,
- $\Omega^{(f)} = \{\mathbf{m}_{l_1}, \dots, \mathbf{m}_{l_k}\}$ , composé de  $k \leq n$  points d'intérêt, auquel est associé  $\Gamma^{(f)} = \{c_{l_1}, \dots, c_{l_k}\}$ .

Alors :

$$\begin{cases} \Omega_0^{(f)} = \emptyset \\ \Omega_j^{(f)} = \Omega_{j-1}^{(f)} \cup \{\mathbf{m}_{l_j}\}, \forall j \in [1, k]_{\mathbb{N}}; \\ \Gamma_0^{(f)} = \emptyset \\ \Gamma_j^{(f)} = \Gamma_{j-1}^{(f)} \cup \{c_{l_j}\}, \forall j \in [1, k]_{\mathbb{N}} \end{cases} \quad (3.131)$$

où :

$$c_{l_j} = \begin{cases} \text{argmax} & \sum_{q=1}^h \left( F_{SFS}^{\sigma(q),j}(c) \right)^2 \\ c \in \{c_{l_{q,j}}\}_{q \in [1,p]_{\mathbb{N}}} \end{cases}, \quad (3.132)$$

où :

- \*  $\forall q \in [1, p-1]_{\mathbb{N}}, F_{SFS}^{\sigma(q),j}(c) \geq F_{SFS}^{\sigma(q+1),j}(c)$ ,
- \*  $\sigma \in \Sigma_p$  est une permutation sur  $[1, p]_{\mathbb{N}}$ ,
- \*  $h = \frac{\alpha}{100}p$ .

$$c_{l_{q,j}} = \begin{cases} \text{argmax} & F_{SFS}^{q,j}(c) \\ c \in \Gamma^{(i)} \setminus \Gamma_{j-1}^{(f)} \end{cases}, \quad (3.133)$$

$$F_{SFS}^{q,j}(c) = \text{Det} \left( \frac{1}{\sigma_q^2} \begin{bmatrix} A^{j-1} & c \end{bmatrix}^T \begin{bmatrix} A^{j-1} & c \end{bmatrix} \right), \quad (3.134)$$

$$\sigma_q^2 = \frac{1}{2} \frac{\| \begin{bmatrix} A^{j-1} & c \end{bmatrix} \hat{\beta} - b_q^T \|_2^2}{m-j}, \quad (3.135)$$

$$\hat{\beta} = \begin{cases} \text{argmin} & \frac{1}{2} \| \begin{bmatrix} A^{j-1} & c \end{bmatrix} \beta - b_q^T \|_2^2 \\ \beta \in \mathbb{R}^j \end{cases}, \quad (3.136)$$

$$\begin{cases} A^0 = [ & ] \\ A^j = \begin{bmatrix} A^{j-1} & c_{l_j} \end{bmatrix} \end{cases}. \quad (3.137)$$

---

### Remarques

1) Typiquement,  $\alpha = 75$ .

2) Une amélioration possible concerne  $\Omega_0^{(f)}$  :

L'utilisation finale des résultats de cette sélection de points d'intérêt est la prédiction, par l'intermédiaire de la relation linéaire «  $\widehat{\Delta\beta} = \hat{A}\Delta\mathbf{I}$  », des  $p$  paramètres du modèle de mouvement. Il paraît donc nécessaire (ou pour le moins prudent) d'envisager uniquement des  $\hat{A}$  de rang maximal  $p$ , donc issus d'au moins  $p$  points d'intérêt. Que

penser, par exemple, d'un prédicteur  $\hat{A} \in \mathbb{R}^{p \times 1}$  (une seule colonne) pour fournir les huit paramètres d'une homographie ? ! Ceci conduit à démarrer avec  $|\Omega_0^{(f)}| \geq p$ . Mais alors quels  $p$  points d'intérêt choisir ? Par exemple, ceux fournis par la procédure présentée dans la section 3.4.1.3.B...

### 3.4.2 Sélection d'un sous-ensemble : approche algébrique

#### 3.4.2.1 Principe informel

La procédure de Golub *et al.* [GKS76, GL96] extrait la sous-matrice  $B$  (d'une matrice  $A$ ) de rang  $k$ , la mieux conditionnée.  $B$  est alors constituée des colonnes les plus linéairement indépendantes donc porteuses du « maximum d'information ». L'algorithme de Golub et Van Loan [[GL96], algorithme 12.2.1, p. 593] est fondé sur une décomposition en valeurs singulières (SVD, en anglais, *Singular Value Decomposition*) de  $A$ . Il exploite le résultat fondamental établi par Eckart et Young [EY36] en 1936 concernant la matrice «  $B$  » de rang  $k$ , la plus proche au sens de la norme de Frobenius de la matrice «  $A$  ».

Cette approche de sélection existe sous deux versions [GKS76]. Tandis que la version de l'algorithme 3.4.1 « ajoute séquentiellement » des colonnes indépendantes, celle de Björck [[Bjö96], §2.7.7, p. 113], décrite dans l'algorithme 3.4.2, « retire séquentiellement » des colonnes dépendantes.

L'algorithme 3.4.1 s'impose pour  $k \leq \frac{n}{2}$  et inversement pour celui de Björck.

#### 3.4.2.2 Algorithme de Golub *et al.*

Soient  $A \in \mathbb{R}^{m \times n}$  et  $b \in \mathbb{R}^m$ , l'algorithme suivant calcule :

- une permutation  $P$ , telle que les  $k$  premières colonnes de  $B = AP$  sont indépendantes,
- un rang  $k$ ,
- un vecteur  $z \in \mathbb{R}^k$  qui est solution du problème de minimisation :

$$\begin{cases} \min & \| B(:, 1:k)z - b \|_2 \\ z \in \mathbb{R}^k \end{cases} . \quad (3.138)$$

- 1) Calculer la SVD de  $A$  :  $A = U\Sigma V^T$ .
- 2) Déterminer  $k \leq \text{rang}(A)$ .
- 3) Appliquer une décomposition QR avec pivotement des colonnes à  $V(:, 1:k)^T$  :

$$Q^T V(:, 1:k)^T P = R = \begin{bmatrix} R_{11} & R_{12} \end{bmatrix}$$

et noter  $AP = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$  avec  $B_1 \in \mathbb{R}^{m \times k}$  et  $B_2 \in \mathbb{R}^{m \times (n-k)}$ .

- 4) Déterminer  $z \in \mathbb{R}^k$  tel que :

$$z = \begin{cases} \operatorname{argmin} & \| b - B_1 z \|_2 \\ z \in \mathbb{R}^k \end{cases} . \quad (3.139)$$

Algorithme 3.4.1: Algorithme de sélection d'un sous-ensemble de Golub et Van Loan.

### 3.4.2.3 Algorithme de Ake Björck

Soient  $A \in \mathbb{R}^{m \times n}$  et  $b \in \mathbb{R}^m$ , l'algorithme suivant calcule :

- une permutation  $P$  telle que les  $k$  premières colonnes de  $AP$  sont suffisamment indépendantes,
- un vecteur  $z \in \mathbb{R}^k$  qui est solution du problème de minimisation :

$$\begin{cases} \min & \| AP \begin{bmatrix} z \\ 0_{n-k,1} \end{bmatrix} - b \|_2, \\ z \in \mathbb{R}^k \end{cases} \quad (3.140)$$

où  $0_{m,n}$  est la matrice nulle de  $m$  lignes et  $n$  colonnes.

- 1) Calculer la SVD de  $A$  :  $A = U\Sigma V^T$ .
- 2) Déterminer  $k \leq \text{rang}(A)$ .
- 3) Soit :

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \text{ où } V_{11} \in \mathbb{R}^{k \times k} \text{ et } V_{22} \in \mathbb{R}^{(n-k) \times (n-k)}.$$

Appliquer une décomposition QR avec pivotement des colonnes à  $W = \begin{bmatrix} V_{22}^T & V_{12}^T \end{bmatrix}$  :

$$Q^T W \tilde{P} = R.$$

- 4) Soit  $P_I$  la matrice de permutation telle que :

$$\begin{bmatrix} V_{12}^T & V_{22}^T \end{bmatrix} P_I = \begin{bmatrix} V_{22}^T & V_{12}^T \end{bmatrix}, \text{ c'est-à-dire } P_I = \begin{bmatrix} 0_{n-k,k} & Id_{n-k} \\ Id_k & 0_{k,n-k} \end{bmatrix},$$

où  $Id$  est la matrice identité.

- 5) Soit  $AP = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$ , où  $P = P_I^T \tilde{P} P_I$  et  $B_1 \in \mathbb{R}^{m \times k}$ . Calculer la décomposition QR de  $B_1$  et résoudre :

$$z = \begin{cases} \text{argmin} & \| B_1 z - b \|_2. \\ z \in \mathbb{R}^k \end{cases} \quad (3.141)$$

Algorithme 3.4.2: Algorithme de sélection d'un sous-ensemble de Ake Björck.

## 3.5 Sélection pour le suivi par prédiction linéaire

### 3.5.1 La méthode de suivi

La méthode de suivi qui nous intéresse est celle de Jurie et Dhome [JD02]. Elle a été présentée en détail à la section 2.5.1. Elle est fondée sur l'apprentissage d'un prédicteur linéaire  $\hat{A}$  qui fait intervenir un ensemble de points  $\Omega^{(f)}$ .

### 3.5.2 Le problème

La question que nous nous posons est donc de déterminer le meilleur ensemble de points  $\Omega^{(f)}$  afin d'obtenir un meilleur prédicteur et donc de suivre mieux et plus longtemps la région d'intérêt. Dans la suite, nous allons donner quelques réponses à cette question.

### 3.5.3 L'algorithme général de sélection de points d'intérêt

Les grandes lignes de notre algorithme de « sélection-validation » sont décrites dans l'algorithme 3.5.1 et chaque étape est ensuite détaillée.

- 1) Déterminer  $\Omega^{(i)}$ , un ensemble initial de points - - cf. §3.6.1
- 2) Sélectionner un sous-ensemble  $\Omega^{(f)}$  de  $m$  points appartenant à  $\Omega^{(i)}$  - - cf. §3.6.2
- 3) Estimer, à partir de  $\Omega^{(f)}$ , le prédicteur  $A$  du modèle linéaire sous-jacent à la relation  $\Delta\beta = A\Delta\mathbf{I}$  - - cf. §2.5.1
- 4) Valider  $\Omega^{(f)}$  à l'aide d'une mesure  $EV$  (erreur de validation) - - cf. §3.7.2.5

Algorithme 3.5.1: Les grandes lignes de notre algorithme.

---

#### Remarque

Les comparaisons sont faites en fixant le nombre de points d'intérêt.

---

## 3.6 Implémentation des algorithmes de sélection des points d'intérêt

### 3.6.1 Détermination de l'ensemble initial

L'ensemble initial  $\Omega^{(i)}$  de points  $\bar{X}$ , ici dénommés  $\mathbf{m}$ , est obtenu en appliquant à l'image un détecteur de points de Moravec (voir §3.2.2), de Harris et Stephens (voir §3.2.4), de Shi et Tomasi (KLT) (voir §3.2.5), de Smith et Brady (SUSAN) (voir §3.2.6) ou de Lowe (SIFT) (voir §3.2.8).

### 3.6.2 Sélection d'un sous-ensemble

Le sous-ensemble  $\Omega^{(f)}$  est obtenu en appliquant séquentiellement un ou plusieurs filtres de sélection : algorithme 3.6.1.

Initialisations :  
 $ens \leftarrow \Omega^{(i)}$  - - l'ensemble initial de points  
 $filtres \leftarrow$  les nbf filtres successifs à appliquer

Algorithme :  
**Pour**  $i$  **de** 1 **à**  $nbf$  **Faire**  
     $ens \leftarrow appliquer\_filtre(filtres(i), ens)$   
**Fin Pour**  
 $\Omega^{(f)} \leftarrow ens$

Algorithme 3.6.1: L'algorithme de sélection d'un sous-ensemble de points en appliquant séquentiellement plusieurs filtres.

Les cinq filtres que nous avons étudiés sont :

- la sélection des meilleurs points selon l'ordre induit par le critère de détection associé à  $\Omega^{(i)} : \mathcal{M}$  (équ. (3.18)),  $\mathcal{HS}$  (équ. (3.53)),  $\mathcal{ST}$  (équ. (3.81)),  $\mathcal{SB}$  (équ. (3.83)) ou  $\mathcal{L}$  (équ. (3.100)) ;
- la sélection d'un sous-ensemble suivant l'approche algébrique de Golub *et al.* (voir §3.4.2). Notre contribution personnelle consiste « tout simplement » à faire «  $A$  » =  $C^T$  (équation (2.108)) dans l'algorithme 3.4.1. En effet, sélectionner une colonne de  $C^T$  équivaut à sélectionner un point d'intérêt étant donné que celle-ci provient des mesures effectuées sur ce point ;
- la sélection d'un sous-ensemble suivant l'approche statistique : se reporter à la section 3.4.1.4 ;
- la sélection des meilleurs points au sens de Zivkovic et Van der Heijden (voir §3.3.2) ;
- la sélection par compartiments. Les points sont sélectionnés, selon un des filtres précédents, en nombre égal dans neuf compartiments identiques qui divisent la région d'intérêt.

### 3.7 Expérimentation : la sélection de points d'intérêt et la validation des prédictors

#### 3.7.1 Les données

Les expériences ont été menées sur cinq images (de dimensions «  $m$  »  $\times$  «  $n$  ») correspondant à des scènes diverses, à savoir :

Nom	$m \times n$	Figure	Accès
Artichaut (Art.)	480 $\times$ 512	3.13-a	banque CMU <a href="http://vasc.ri.cmu.edu/idb/html/motion/artichoke/index.html">http://vasc.ri.cmu.edu/idb/html/motion/artichoke/index.html</a> et [ST94]
Faugeras (Fau.)	423 $\times$ 463	3.13-b	<a href="http://www.lmet.fr/icons/Scans13/Big/9780/26/20/61/582.gif">http://www.lmet.fr/icons/Scans13/Big/9780/26/20/61/582.gif</a>
Cameraman (Cam.)	256 $\times$ 256	3.13-c	chemin d'accès MATLAB <a href=".../toolbox/images/indemos/cameraman.tif">.../toolbox/images/indemos/cameraman.tif</a>
Minneapolis (Min.)	193 $\times$ 227	3.13-d	<a href="http://cvpr.cv.ri.cmu.edu/i/sky-home-L.jpg">http://cvpr.cv.ri.cmu.edu/i/sky-home-L.jpg</a>
Noé (Noé)	240 $\times$ 320	3.13-e	<a href="http://www.enseeiht.fr/~charvi/noe.tif">http://www.enseeiht.fr/~charvi/noe.tif</a>

#### 3.7.2 Protocole

##### 3.7.2.1 Les ensembles initiaux

Ils sont formés, d'au plus 1000 points, provenant des :

- **coins de Harris** (H) (figure 3.14-a) : nous utilisons le code de la boîte à outils de MATLAB « *Machine Vision Toolbox for Matlab* » (<http://www.cat.csiro.au/ict/staff/pic/vision/>) pour lequel nous avons paramétré la taille de la fenêtre des minima locaux (5  $\times$  5) et les dimensions (7  $\times$  7) et variance ( $\sigma = 1$ ) du masque gaussien ;
- **coins de Moravec** (M) (figure 3.14-b) : implémentation « maison » ;



(a)



(b)



(c)



(d)



(e)

FIG. 3.13 – Les cinq images de référence : (a) « Artichaut », (b) « Faugeras », (c) « Cameraman », (d) « Minneapolis » et (e) « Noé ».

- **points de Shi et Tomasi (KLT)** (figure 3.14-c) : nous utilisons le code fourni par Birchfield que nous adaptons pour faire une recherche de maxima locaux en partant de la valeur (nombre réel) de la plus petite des valeurs propres ;
- **coins de Smith et Brady (SUSAN)** (Su) (figure 3.14-d) : implémentation « maison » ;
- **points de Lowe (SIFT)** (figure 3.14-e) : nous utilisons :
  - \* le code exécutable fourni par D. Lowe accessible à l'adresse : <http://www.cs.ubc.ca/~lowe/keypoints/>,
  - \* l'implémentation de Vedaldi (que nous a recommandée D. Lowe accessible à l'adresse <http://vision.ucla.edu/~vedaldi/code/sift/sift.html>) (sources en C) que nous avons adaptée pour recueillir en sortie le critère  $\mathcal{L}$  (équ. (3.100)).

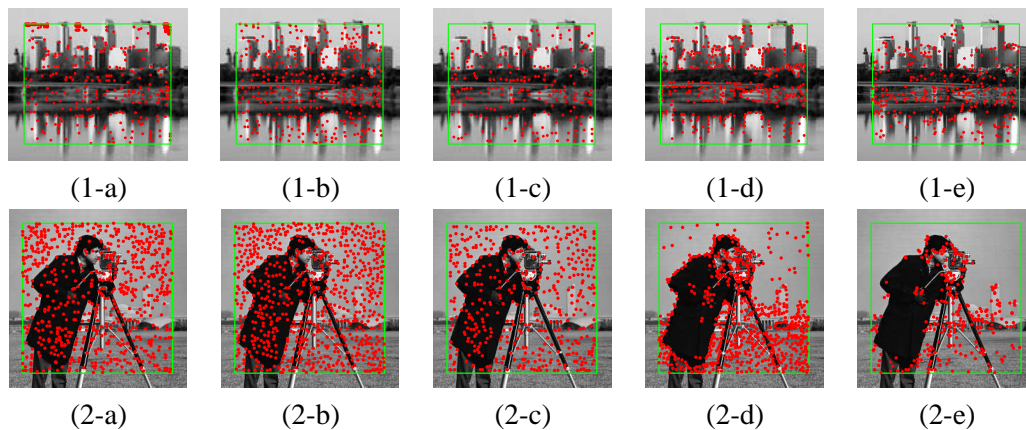


FIG. 3.14 – Deux des cinq images de test et quelques exemples d'ensembles initiaux  $\Omega^{(i)}$  de points d'intérêt (en rouge). Première ligne = « Minneapolis » ; Deuxième ligne = « Cameraman ». Tous les points de : (a) Moravec, (b) Harris, (c) KLT, (d) SUSAN, (e) SIFT.

### 3.7.2.2 Les filtres

À partir des  $\Omega^{(i)}$ , nous sélectionnons un sous-ensemble de points en enchaînant un ou plusieurs des filtres suivants :

- la sélection des meilleurs selon l'ordre induit par le critère associé à  $\Omega^{(i)}$  (B),
- la sélection d'un sous-ensemble proposée par Golub *et al.* (S),
- la sélection séquentielle par ajout (SFS),
- le critère de Zivkovic et Van der Heijden (ZV) (nous utilisons le code fourni par Zoran Zivkovic : <http://staff.science.uva.nl/~zivkovic/Publications/codebf.zip>),
- la sélection par compartiments (C). La zone d'intérêt est divisée en neuf compartiments (figures 3.15-(3-d), (4-d) et (5-d)).

Nous pouvons ainsi facilement identifier nos tests. L'ordre des filtres se lit de gauche à droite ; ainsi par exemple :

- H\_B signifie que nous avons choisi les  $m$  meilleurs points de Harris,
- KLT\_S\_C signifie que nous avons choisi la sélection d'un sous-ensemble définie par Golub *et al.*, par compartiments, en partant de l'ensemble des points de KLT.

La figure 3.15 illustre quelques  $\Omega^{(f)}$  obtenus par différentes cascades de filtres.



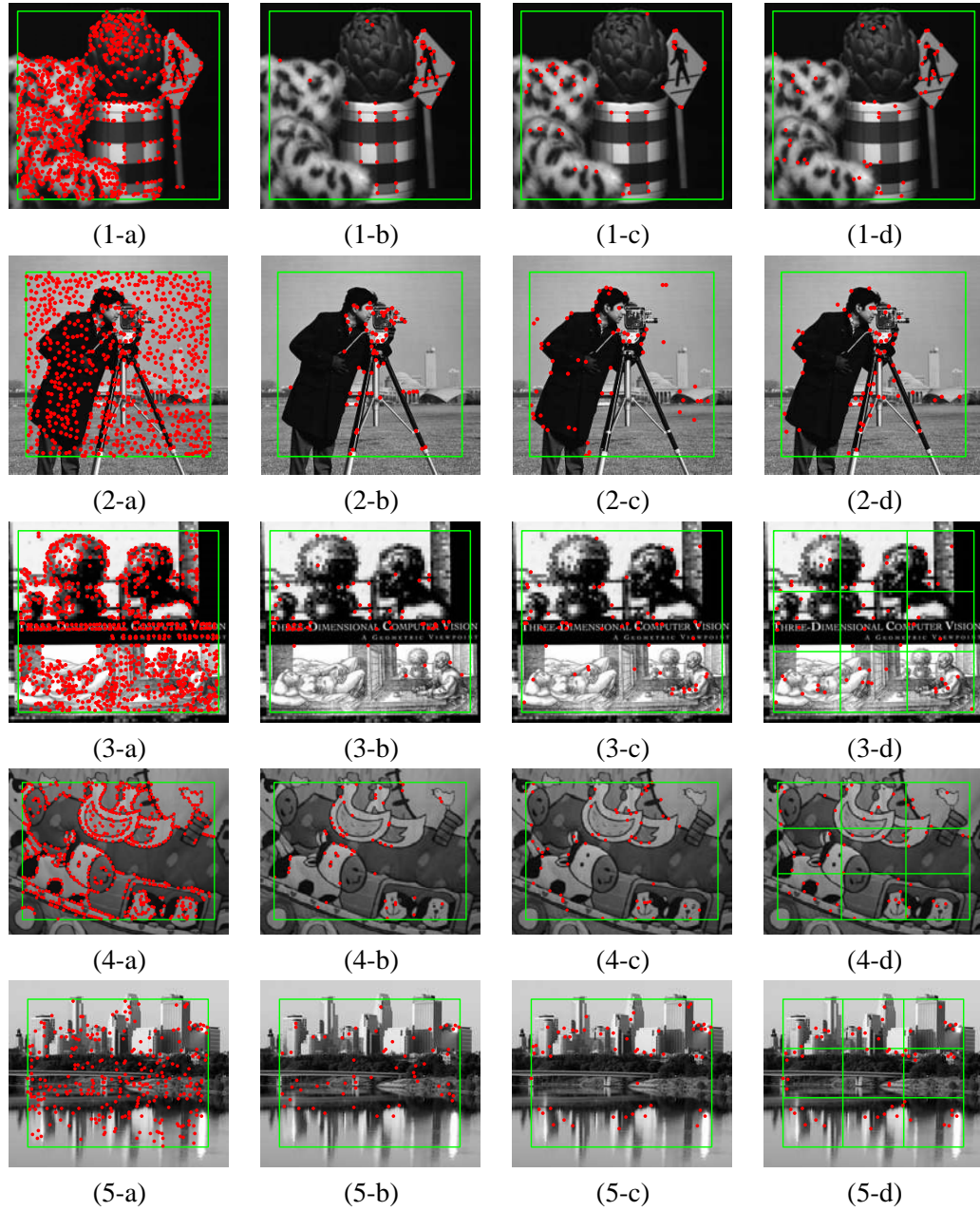


FIG. 3.15 – Les cinq images de test et quelques exemples d'ensembles initiaux  $\Omega^{(i)}$  et de sous-ensembles de points d'intérêt  $\Omega^{(f)}$  (en rouge) pour  $m = 60$  points et  $a = 15$  pixels. Première ligne = « Artichaut » : (1-a) tous les points de Moravec ; 60 points obtenus par : (1-b) M\_B, (1-c) M\_ZV et (1-d) M\_S. Deuxième ligne = « Cameraman » : (2-a) tous les points de Harris ; 60 points obtenus par : (2-b) H\_B, (2-c) H\_ZV et (2-d) H\_S. Troisième ligne = « Faugeras » : (3-a) tous les points de KLT ; 60 points obtenus par : (3-b) KLT\_B, (3-c) KLT\_S et (3-d) KLT\_S\_C. Quatrième ligne = « Noé » : (4-a) tous les points de SUSAN ; 60 points obtenus par : (4-b) Su\_B, (4-c) Su\_S et (4-d) Su\_S\_C. Cinquième ligne = « Minneapolis » : (5-a) tous les points de SIFT ; 60 points obtenus par : (5-b) SIFT\_B, (5-c) SIFT\_SFS et (5-d) SIFT\_SFS\_C.

### 3.7.2.3 Les déformations

La déformation  $T$  est une homographie paramétrée par les déplacements des sommets du quadrilatère d'intérêt  $\mathcal{R}$ .

Pendant l'étape d'apprentissage, les variations des huit ( $4 \times 2$ ) paramètres  $\beta$  sont prises aléatoirement et uniformément pour chaque paramètre dans l'intervalle  $[-a, a]$ , où  $a$  représente l'amplitude maximale de la perturbation :

$$\forall i \in [1, 8]_{\mathbb{N}}, \beta_i \sim \mathcal{U}([-a, a]).$$

Typiquement,  $a \in J_a = \{1, 3, 5, 10, 15\}$ , où l'unité est le pixel.

Pour chaque test, nous utilisons le même ensemble d'apprentissage  $\mathcal{A}$  ( $N_a = 1500$ ) et le même ensemble de tests  $\mathcal{T}$  ( $N_t = 1000$ ).

Le nombre  $m$  de points finalement sélectionnés est tel que :

$$m = |\Omega^{(f)}| \in J_m = \{20, 40, 60, 80, 100, 120\}.$$

---

#### Remarque

Lorsque le modèle de déformations  $T$  est une translation, si :

- les coordonnées des points de  $\Omega^{(f)}$  ont les mêmes parties fractionnaires (par exemple nulles),
- les niveaux de gris aux points de coordonnées non entières sont calculés par interpolation bilinéaire à partir des quatre pixels voisins (cf. remarque de la section 4.1.3.1.2),

alors le rang  $r$  de la matrice  $C$  intervenant lors de l'estimation du prédicteur linéaire  $\hat{A}$  (équation (2.106)) peut être très inférieur à  $\min(m, N_a)$ .  $r$  dépend du nombre (indépendant des points  $\mathbf{m}$ ) de pixels impliqués dans le calcul de l'ensemble des niveaux de gris  $\{I(T(\mathbf{m}, \Delta\beta^l))\}_{l \in [1, N_a]_{\mathbb{N}}}$ . Par exemple, si  $\Delta\beta^l \in B_{\infty}^f(O, a)$ , où  $a \in \mathbb{N}^*$  est l'amplitude du déplacement de la translation globale, alors :

$$r \leq (2a + 1)^2.$$

En effet, une ligne  $l \in [1, N_a]_{\mathbb{N}}$  de la matrice  $C^T \in \mathbb{R}^{N_a \times m}$  peut s'écrire :

$$\begin{aligned} C_l^T &= [I(T(\mathbf{m}_1, \Delta\beta^l)) - I(\mathbf{m}_1) \quad \cdots \quad I(T(\mathbf{m}_m, \Delta\beta^l)) - I(\mathbf{m}_m)] \in \mathbb{R}^m \\ &= \left[ \sum_{i=1}^4 \alpha_i^l (I(V(T(\mathbf{m}_1, \Delta\beta^l), i)) - I(\mathbf{m}_1)) \quad \cdots \quad \sum_{i=1}^4 \alpha_i^l (I(V(T(\mathbf{m}_m, \Delta\beta^l), i)) - I(\mathbf{m}_m)) \right] \\ &= \underbrace{\begin{bmatrix} \alpha_1^l & \cdots & \alpha_4^l \end{bmatrix}}_{\tilde{A}} \underbrace{\begin{bmatrix} I(V(T(\mathbf{m}_1, \Delta\beta^l), 1)) - I(\mathbf{m}_1) & \cdots & I(V(T(\mathbf{m}_m, \Delta\beta^l), 1)) - I(\mathbf{m}_m) \\ \vdots & & \vdots \\ I(V(T(\mathbf{m}_1, \Delta\beta^l), 4)) - I(\mathbf{m}_1) & \cdots & I(V(T(\mathbf{m}_m, \Delta\beta^l), 4)) - I(\mathbf{m}_m) \end{bmatrix}}_{\tilde{B}} \end{aligned} \quad (3.142)$$

où  $\forall i \in [1, 4]_{\mathbb{N}}$  :

- $\alpha_i^l$  est le  $i$ -ème coefficient de l'interpolation bilinéaire ;
- $V(T(\mathbf{m}, \Delta\beta^l), i)$  est le  $i$ -ème pixel parmi les quatre plus proches (au sens de la norme euclidienne) voisins de  $T(\mathbf{m}, \Delta\beta^l)$ . Regroupons tous les points de l'ensemble  $\{T(\mathbf{m}, \Delta\beta^l)\}_{l \in [1, N_a]_{\mathbb{N}}}$  qui partagent les mêmes quatre plus proches voisins.

Alors,  $C^T$  peut s'écrire  $C^T = AB$ , où :

- $A$  contient les coefficients de l'interpolation bilinéaire :

$$A = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_K \end{bmatrix} \in \mathbb{R}^{N_a \times 4K},$$

où :

$$\forall j \in [1, K]_{\mathbb{N}}, A_j \in \mathbb{R}^{k_j \times 4} = [\alpha_{j'}^{i'}]_{i' \in [\sum_{i=0}^{j-1} k_i + 1, \sum_{i=0}^j k_i]_{\mathbb{N}}, j' \in [1, 4]_{\mathbb{N}}},$$

avec  $k_0 = 0$  et  $\sum_{j=1}^K k_j = N_a$ .

- $B$  contient les différences de niveaux de gris :

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_K \end{bmatrix} \in \mathbb{R}^{4K \times m},$$

où  $\forall j \in [1, K]_{\mathbb{N}}, B_j \in \mathbb{R}^{4 \times m}$ . Par exemple, en reprenant l'équation (3.142), si  $l \in [k_1 + 1, k_1 + k_2]_{\mathbb{N}}$ , alors  $B_2 = \tilde{B}$ .

De cette décomposition, découlent :

-

$$\begin{cases} \forall j \in [1, K]_{\mathbb{N}}, rg(A_j) \leq \min(4, k_j) \\ rg(A) \leq \sum_{j=1}^K rg(A_j) \end{cases} \Rightarrow rg(A) \leq 4K;$$

- une condition suffisante pour que deux lignes de  $B$  soient identiques est qu'elles impliquent le même pixel. Le rang de  $B$  est donc majoré par le nombre  $L$  de pixels différents sous-jacents :

$$rg(B) \leq L \leq 4K;$$

-

$$r = \min(rg(A), rg(B)) \leq L.$$

Le rang maximal de  $C$  étant connu, il en est de même pour :

- le nombre maximal de points,
- le nombre maximal de perturbations.

### 3.7.2.4 Les séquences artificielles

L'algorithme de suivi utilisé, qui reprend les idées de Cootes *et al.* [CET01] et de Jurie et Dhome (JD) [JD02], est implémenté par Matthijs Douze [Dou04] et offre de plus la possibilité (que nous exploitons) de faire du suivi hiérarchique. La hiérarchisation consiste à utiliser séquentiellement des prédicteurs  $A$ , adaptés à une amplitude de plus en plus faible. À chaque niveau de la hiérarchie, l'algorithme JD est appliqué pour l'amplitude donnée. Ainsi, alors que S15 signifie que l'algorithme JD est utilisé pour la seule amplitude  $a = 15$  pixels, S15-10-5 indique un suivi hiérarchique d'amplitudes successives 15, 10 et 5 pixels.

Chacune des cinq images  $I_{ref}(= I_0)$  donne naissance à des séquences artificielles, pour lesquelles  $\forall i \geq 0, I_{i+1}$  provient de  $I_i$  par application d'une homographie d'amplitude  $a_{i+1}$ , satisfaisant à  $a_{i+1} \geq a_i$ . La transformation homographique est définie globalement

par une rotation suivie d'une perturbation aléatoire de chacun des sommets du quadrilatère. Cette dernière suit une loi uniforme sur  $[-5, 5]$  (pixels). Entre deux images consécutives  $I_i$  et  $I_{i+1}$ , l'angle de rotation est de  $\frac{\pi}{20 \times 180}(i + 1)$  radians et le centre de rotation est le centre de l'image. La figure 3.16 présente pour une séquence artificielle issue de l'image « Noé », trois couples d'images consécutives impliquées dans l'échec du suivi en relation avec  $\Omega^{(f)}$  et la hiérarchie du suivi :

images	$\Omega^{(f)}$	hiérarchie
(c,d)	H_B	S15
(e,f)	M_SFS	S15
(g,h)	M_SFS	S15-10-5-3-1

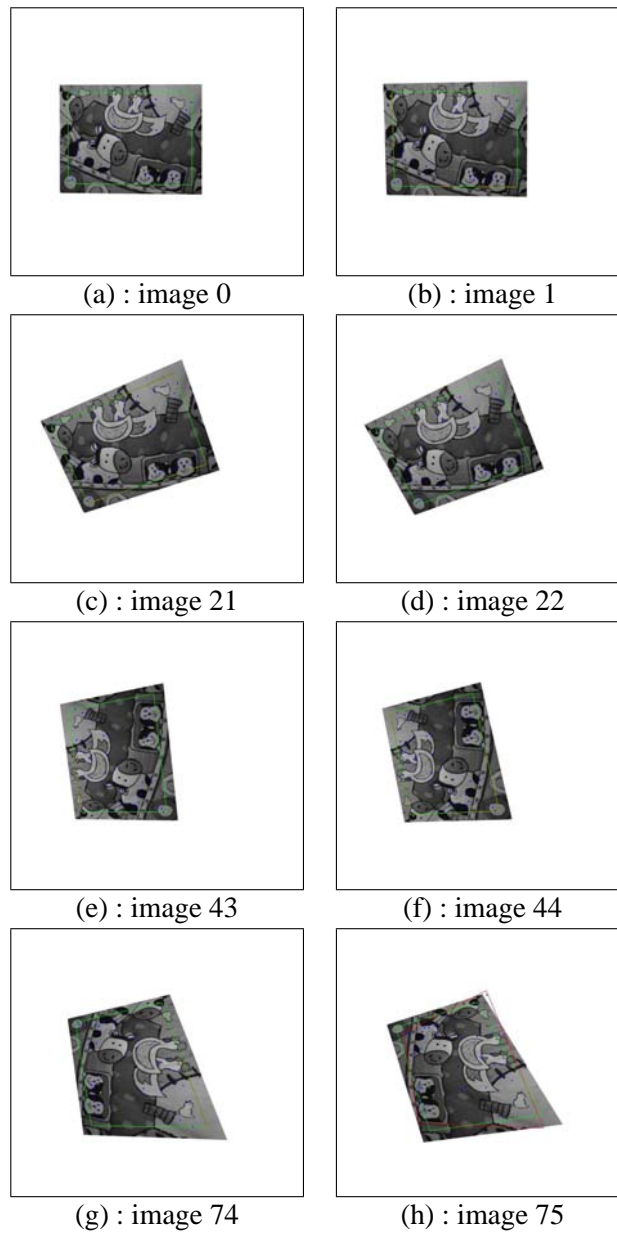


FIG. 3.16 – Séquence artificielle générée à partir de l'image « Noé ».

La figure 3.17 montre l'évolution de l'amplitude maximale des déplacements des sommets du quadrilatère en fonction du numéro de l'image. Ici, à partir de l'image 49, l'amplitude maximale des déplacements à suivre est supérieure à celle des déplacements appris : l'algorithme de suivi est donc voué à échouer.

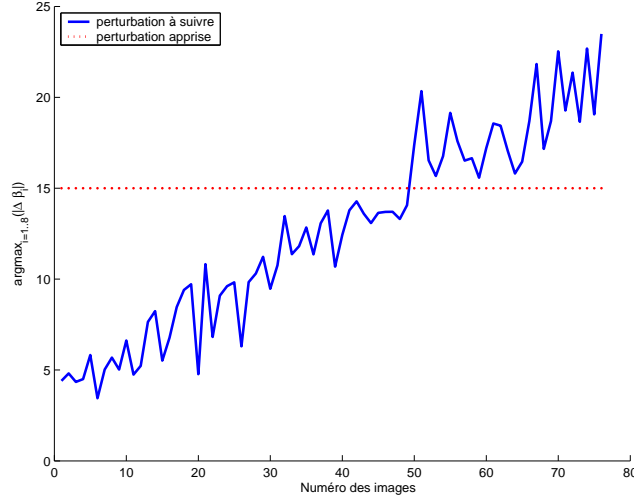


FIG. 3.17 – Évolution de l'amplitude maximale des perturbations à suivre pendant le suivi sur la séquence artificielle présentée sur la figure 3.16.

### 3.7.2.5 Validation

#### A) Le critère local

En rappelant que l'homographie est paramétrée par les déplacements des sommets du quadrilatère d'intérêt  $\mathcal{R}$ , notre jauge associée au prédicteur  $A$  (mesure de qualité), nommée « erreur de validation » EV, est définie, en munissant l'espace produit  $((\mathbb{R}^2)^4)^{N_t}$  de la norme  $\frac{1}{N_t} \|\cdot\|_1$ , par :

$$EV = \frac{1}{N_t} \sum_{j=1}^{N_t} \frac{1}{4} \|\Delta \beta^j - \widehat{\Delta \beta^j}\|_1, \quad (3.143)$$

où :

- $N_t = |\mathcal{T}| = \{|\Delta \beta^j|\}_{j \in [1, N_t]_{\mathbb{N}}}$ ,
- $\Delta \beta^j$  est la  $j$ -ième perturbation « tirée aléatoirement » (donc connue),
- $\widehat{\Delta \beta^j} = A \Delta \mathbf{I}^j$ ,
- $\Delta \mathbf{I}^j = \mathbf{I}_{ref}(\beta_0 + \Delta \beta^j) - \mathbf{I}_{ref}(\beta_0)$ ,
- $(\mathbb{R}^2)^4$  est muni de la norme :  $\frac{1}{4} \|\cdot\|_1$ ,
- $\mathbb{R}^2$  est muni de la norme 2 (euclidienne),
- $\Delta \beta = [(\Delta \beta_1)^T (\Delta \beta_2)^T (\Delta \beta_3)^T (\Delta \beta_4)^T]^T$ ,
- $\forall i \in [1, 4]_{\mathbb{N}}$ ,  $\Delta \beta_i = [\Delta \beta_{ix} \Delta \beta_{iy}]^T$  est le déplacement du sommet  $i$  du quadrilatère d'intérêt  $\mathcal{R}$ .

Cette erreur (EV) est exprimée en pixels.

## B) Le critère sur les séquences

Le procédé de construction des séquences artificielles conduit à un inévitable échec du suivi pour l'image  $I_{i_{max}}$ . On cherche donc  $i_{max}$  qui représente le nombre d'images correctement suivies en appliquant l'algorithme de suivi de Jurie et Dhome implémenté par Matthijs Douze [Dou04]. Ce nombre dépend a priori de  $\Omega^{(f)}$ . Nous comparons donc les  $i_{max}$  associés aux différents  $\Omega^{(f)}$ .

### 3.7.3 Résultats

#### 3.7.3.1 Concernant la validation du prédicteur

##### A) Des chiffres, des courbes

Les figures 3.18, 3.19, 3.20 et 3.21 illustrent une infime partie des résultats obtenus. Chaque graphique donne les erreurs de validation en fonction de  $m = |\Omega^{(f)}|$  pour :

- une image  $I_{ref}$  particulière ;
- un ou plusieurs  $\Omega^{(i)}$  ;
- une amplitude  $a$  ;
- différents  $\Omega^{(f)}$ .

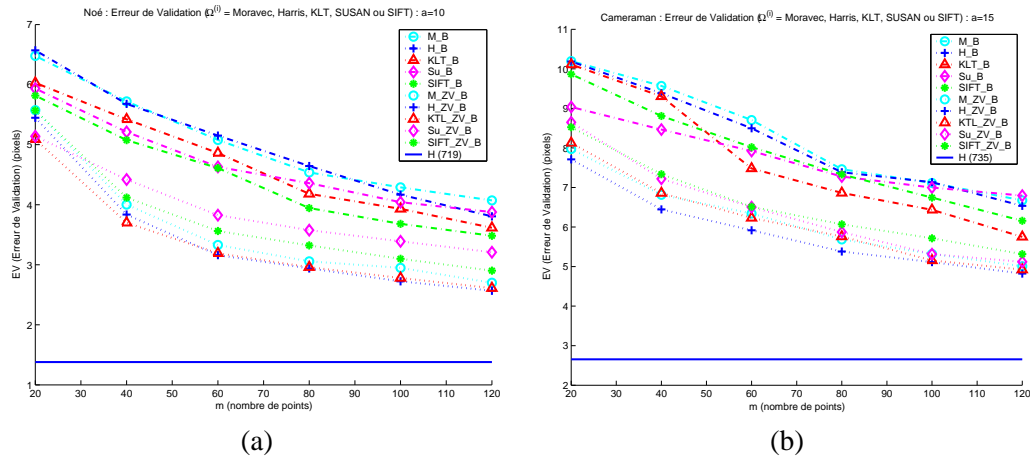


FIG. 3.18 – Erreurs de validation pour différents  $\Omega^{(f)}$  obtenus via les filtres (B) et (ZV). Deux images  $I_{ref}$  « Noé » et « Cameraman ». Deux amplitudes  $a = 10$  ou  $a = 15$  pixels. « L'asymptote » correspond à la plus faible erreur de validation des  $\Omega^{(i)}$  :  $n = |\Omega^{(i)}|$  figure entre parenthèses dans la légende.

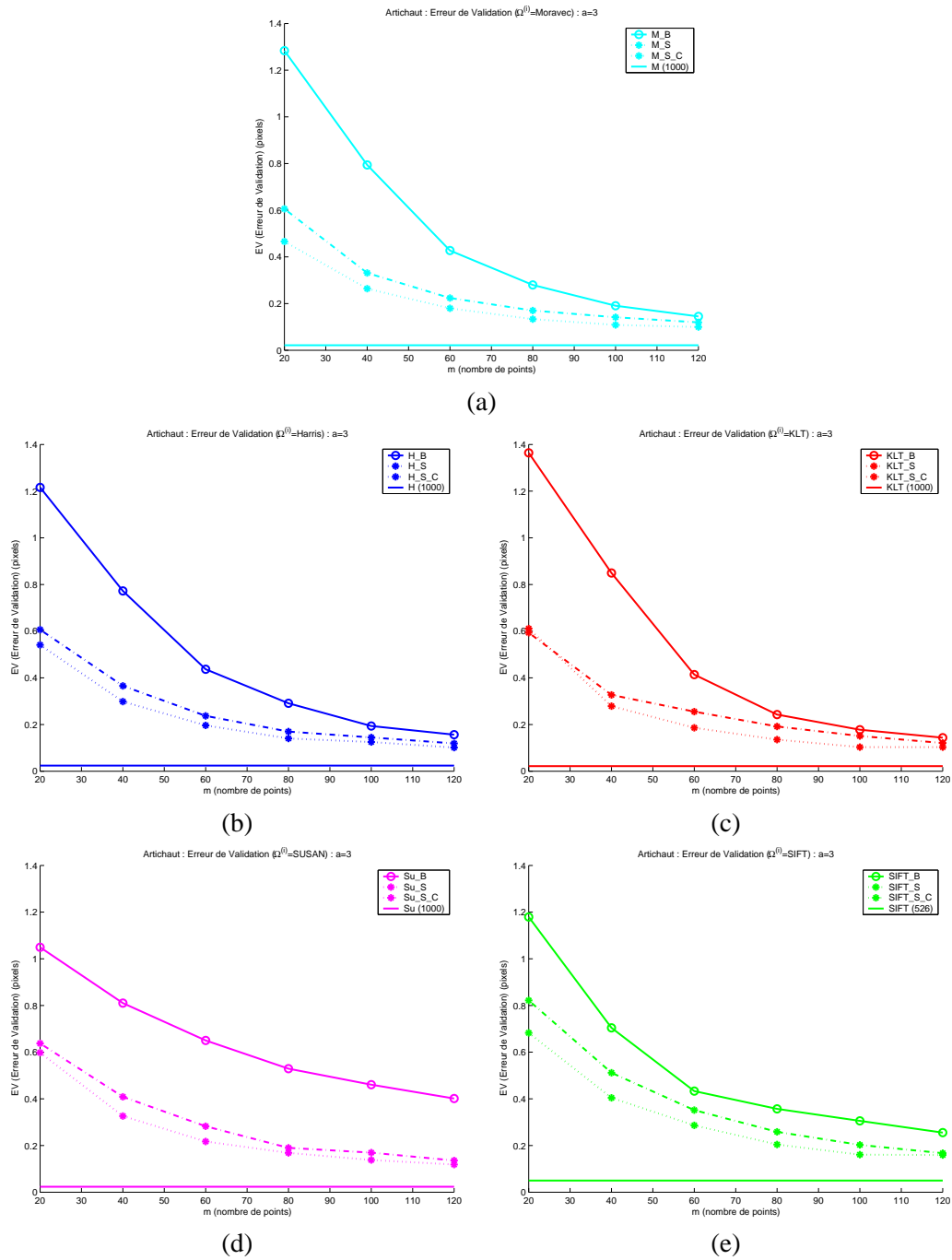


FIG. 3.19 – Erreurs de validation pour différents  $\Omega^{(f)}$  obtenus via les filtres (S) et (C). Une image « Artichaut ». Une amplitude  $a = 3$  pixels. « L'asymptote » correspond à la plus faible erreur de validation des  $\Omega^{(i)}$  :  $n = |\Omega^{(i)}|$  figure entre parenthèses dans la légende.

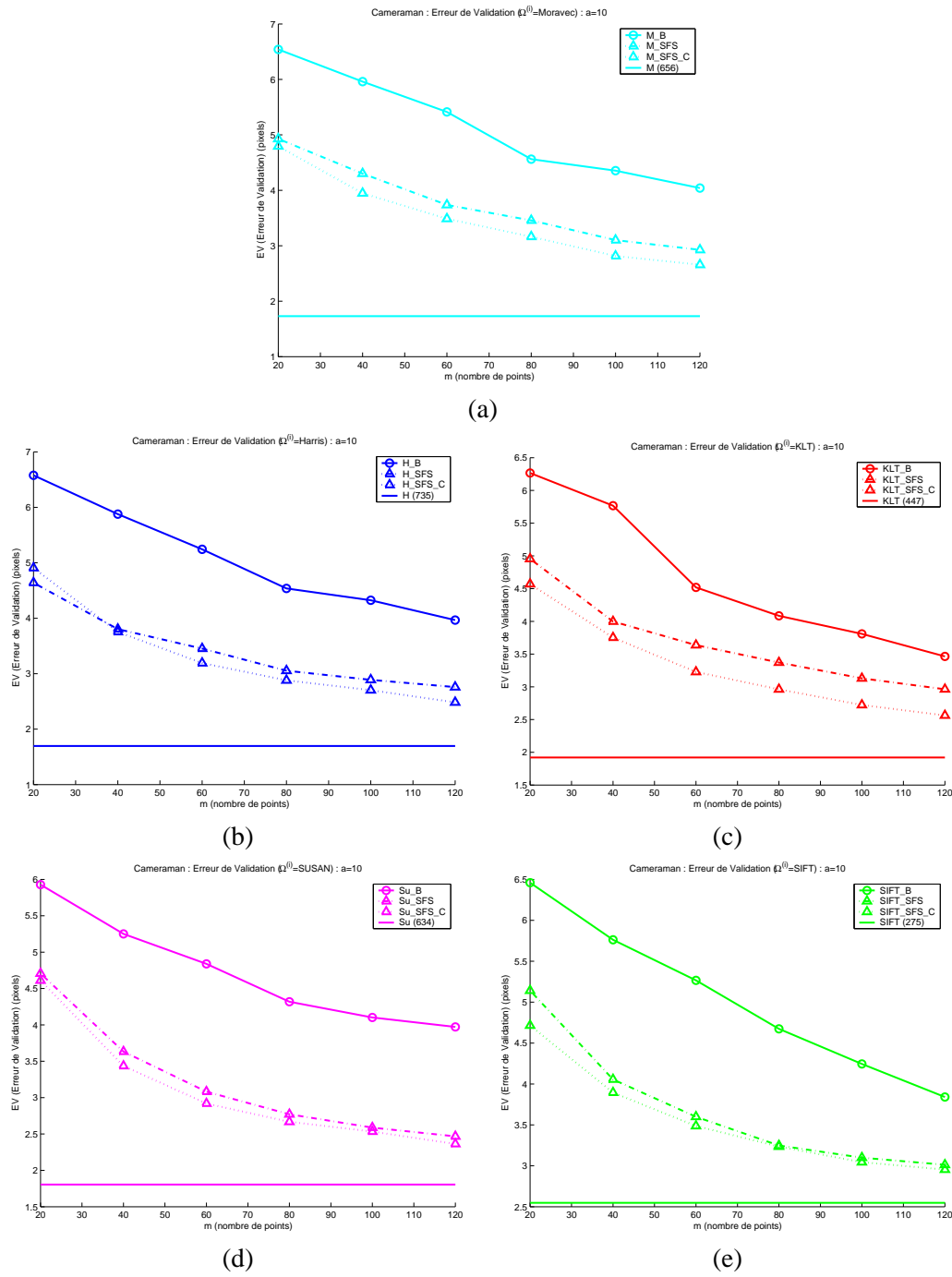


FIG. 3.20 – Erreurs de validation pour différents  $\Omega^{(f)}$  obtenus via les filtres (SFS) et (C). Une image « Cameraman ». Une amplitude  $a = 10$  pixels. « L'asymptote » correspond à la plus faible erreur de validation des  $\Omega^{(i)}$  :  $n = |\Omega^{(i)}|$  figure entre parenthèses dans la légende.



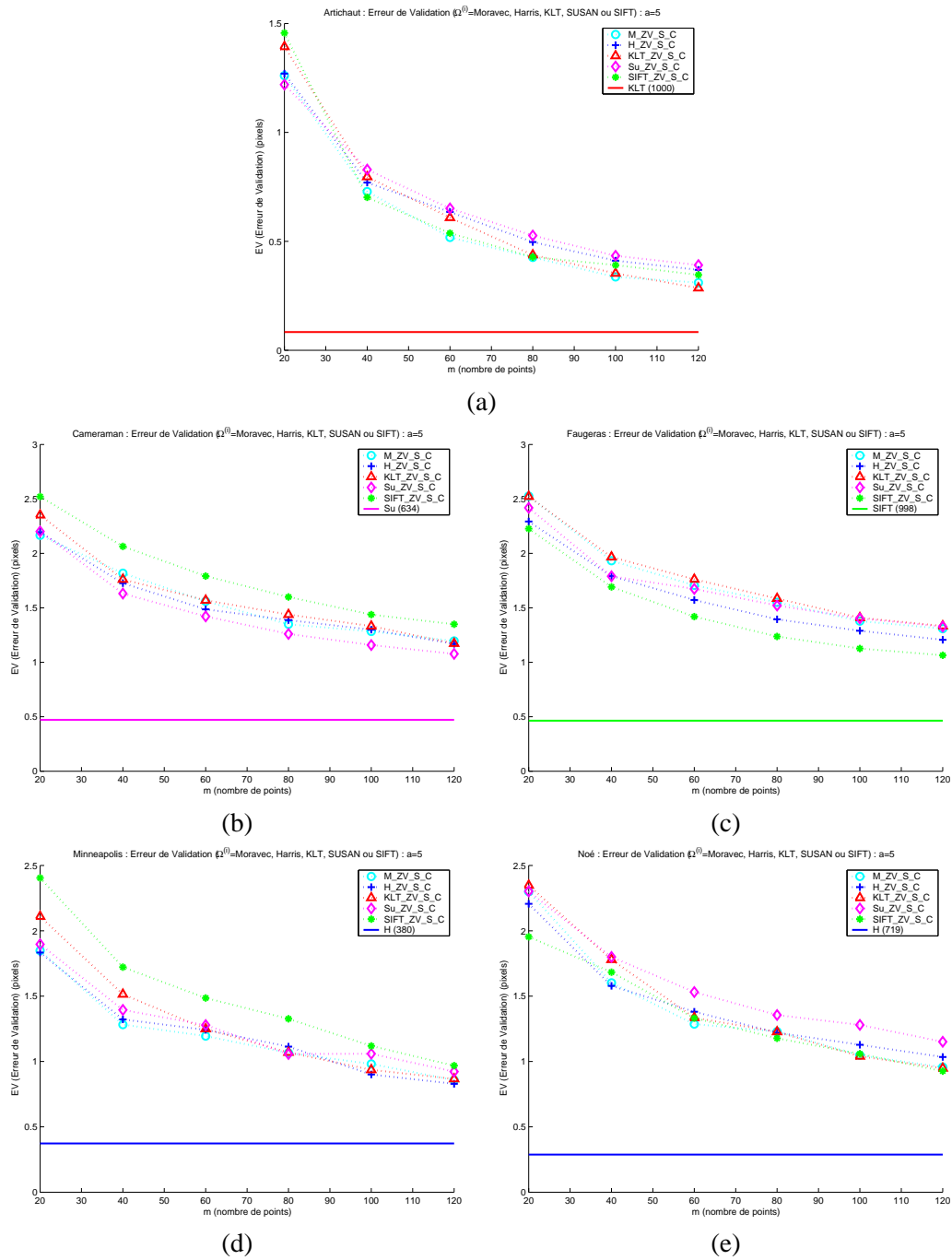


FIG. 3.21 – Erreurs de validation pour différents  $\Omega^{(f)}$ . Différentes images  $I_{ref}$ . Une amplitude  $a = 5$  pixels. « L'asymptote » correspond à la plus faible erreur de validation des  $\Omega^{(i)}$  :  $n = |\Omega^{(i)}|$  figure entre parenthèses dans la légende.

Les tableaux 3.1, 3.2 et 3.3 répertorient, quant à eux, pour chacune des  $I_{ref}$ , les gains moyens minimal (g) et maximal (G) obtenus sur l'erreur de validation.

En rappelant que  $J_a$  représente l'ensemble des amplitudes testées et  $J_m$  l'ensemble des nombres de points d'intérêt de  $\Omega^{(f)}$  testés, l'application  $g_{\Omega^{(i)}\_B \rightarrow \Omega^{(f)}}(a, m)$  donne le gain que l'on a à utiliser l'ensemble  $\Omega^{(f)}$  à la place de l'ensemble  $\Omega^{(i)}\_B$ .

Ainsi, pour chaque  $\Omega^{(i)}$  et chaque  $\Omega^{(f)}$  retenus, à partir des 30 ( $= 6 \times 5$ ) gains :

$$g_{\Omega^{(i)}\_B \rightarrow \Omega^{(f)}}(a, m), \forall a \in J_a, \forall m \in J_m,$$

nous définissons :

- les gains moyens  $\bar{g}(a)$  :

$$\bar{g}(a) = \frac{1}{6} \sum_{m \in J_m} g_{\Omega^{(i)}\_B \rightarrow \Omega^{(f)}}(a, m), \forall a \in J_a,$$

- le gain minimal g :

$$g = \left\{ \min_{a \in J_a} \bar{g}(a) \right\},$$

- le gain maximal G :

$$G = \left\{ \max_{a \in J_a} \bar{g}(a) \right\}.$$

La notation « H\_B  $\rightarrow$  H\_S » indique que le gain moyen est calculé pour passer de l'erreur de validation associée à H\_B à l'erreur de validation associée à H\_S.

$I_{ref}$	$\Omega^{(i)}\_B \rightarrow \Omega^{(f)}$	g (%)   (a)	G (%)   (a)
Cam.	M_B $\rightarrow$ M_ZV_B	24.5 (1)	27.7 (10)
	H_B $\rightarrow$ H_ZV_B	21.6 (1)	27.9 (15)
	KLT_B $\rightarrow$ KLT_ZV_B	-3.9 (1)	19.4 (10)
	Su_B $\rightarrow$ Su_ZV_B	17.5 (15)	30.3 (1)
	SIFT_B $\rightarrow$ SIFT_ZV_B	5.6 (1)	22.0 (5)
Art.	M_B $\rightarrow$ M_ZV_B	21.3 (15)	34.6 (1)
	H_B $\rightarrow$ H_ZV_B	25.3 (15)	35.3 (1)
	KLT_B $\rightarrow$ KLT_ZV_B	-1.4 (1)	27.0 (5)
	Su_B $\rightarrow$ Su_ZV_B	30.2 (15)	51.2 (1)
	SIFT_B $\rightarrow$ SIFT_ZV_B	-9.3 (5)	4.9 (1)
Min.	M_B $\rightarrow$ M_ZV_B	-2.2 (15)	32.6 (1)
	H_B $\rightarrow$ H_ZV_B	4.7 (15)	21.2 (5)
	KLT_B $\rightarrow$ KLT_ZV_B	-13.1 (1)	9.4 (10)
	Su_B $\rightarrow$ Su_ZV_B	2.9 (15)	31.0 (1)
	SIFT_B $\rightarrow$ SIFT_ZV_B	-1.7 (3)	10.4 (10)
Noé	M_B $\rightarrow$ M_ZV_B	19.0 (15)	40.7 (5)
	H_B $\rightarrow$ H_ZV_B	17.1 (1)	32.4 (5)
	KLT_B $\rightarrow$ KLT_ZV_B	14.2 (1)	34.9 (5)
	Su_B $\rightarrow$ Su_ZV_B	11.5 (15)	40.9 (1)
	SIFT_B $\rightarrow$ SIFT_ZV_B	3.6 (3)	15.7 (10)
Fau.	M_B $\rightarrow$ M_ZV_B	28.3 (1)	45.5 (5)
	H_B $\rightarrow$ H_ZV_B	37.7 (15)	68.3 (1)
	KLT_B $\rightarrow$ KLT_ZV_B	19.9 (1)	29.1 (10)
	Su_B $\rightarrow$ Su_ZV_B	33.6 (15)	48.3 (1)
	SIFT_B $\rightarrow$ SIFT_ZV_B	-14.9 (1)	9.5 (10)

TAB. 3.1 – Gains moyens minimal et maximal sur l'erreur de validation pour les cinq images de référence et différents  $\Omega^{(f)}$  obtenus par les filtres (B) et (ZV).

$I_{ref}$	$\Omega^{(i)}\_B \rightarrow \Omega^{(f)}$	g (%)   (a)	G (%)   (a)
Cam.	M_B $\rightarrow$ M_S	13.4 (10)	35.3 (1)
	M_B $\rightarrow$ M_S_C	24.9 (5)	32.5 (1)
	H_B $\rightarrow$ H_S	15.3 (15)	28.0 (1)
	H_B $\rightarrow$ H_S_C	22.1 (3)	29.4 (1)
	KLT_B $\rightarrow$ KLT_S	8.5 (5)	13.2 (3)
	KLT_B $\rightarrow$ KLT_S_C	15.9 (5)	25.6 (15)
	Su_B $\rightarrow$ Su_S	12.0 (15)	55.7 (1)
	Su_B $\rightarrow$ Su_S_C	24.4 (15)	54.2 (1)
	SIFT_B $\rightarrow$ SIFT_S	6.3 (15)	16.4 (5)
	SIFT_B $\rightarrow$ SIFT_S_C	15.6 (3)	23.0 (1)
Art.	M_B $\rightarrow$ M_S	29.6 (15)	40.2 (3)
	M_B $\rightarrow$ M_S_C	38.8 (15)	52.5 (3)
	H_B $\rightarrow$ H_S	31.5 (15)	39.9 (3)
	H_B $\rightarrow$ H_S_C	37.7 (15)	49.1 (3)
	KLT_B $\rightarrow$ KLT_S	20.1 (15)	35.6 (1)
	KLT_B $\rightarrow$ KLT_S_C	31.8 (15)	52.2 (1)
	Su_B $\rightarrow$ Su_S	38.2 (15)	56.5 (3)
	Su_B $\rightarrow$ Su_S_C	44.7 (15)	<b>63.0</b> (3)
	SIFT_B $\rightarrow$ SIFT_S	4.3 (10)	35.2 (1)
	SIFT_B $\rightarrow$ SIFT_S_C	21.9 (10)	41.1 (3)
Min.	M_B $\rightarrow$ M_S	21.7 (15)	32.9 (1)
	M_B $\rightarrow$ M_S_C	19.9 (15)	40.4 (1)
	H_B $\rightarrow$ H_S	20.3 (15)	27.1 (5)
	H_B $\rightarrow$ H_S_C	22.3 (15)	35.7 (1)
	KLT_B $\rightarrow$ KLT_S	10.2 (5)	19.7 (1)
	KLT_B $\rightarrow$ KLT_S_C	16.3 (15)	33.1 (1)
	Su_B $\rightarrow$ Su_S	13.1 (5)	22.7 (1)
	Su_B $\rightarrow$ Su_S_C	16.6 (15)	27.1 (1)
	SIFT_B $\rightarrow$ SIFT_S	6.5 (3)	23.4 (1)
	SIFT_B $\rightarrow$ SIFT_S_C	18.6 (15)	26.8 (1)
Noé	M_B $\rightarrow$ M_S	14.1 (15)	31.4 (1)
	M_B $\rightarrow$ M_S_C	17.4 (15)	30.8 (5)
	H_B $\rightarrow$ H_S	12.8 (1)	20.1 (5)
	H_B $\rightarrow$ H_S_C	18.1 (15)	24.7 (5)
	KLT_B $\rightarrow$ KLT_S	10.9 (15)	31.5 (1)
	KLT_B $\rightarrow$ KLT_S_C	13.4 (3)	30.5 (1)
	Su_B $\rightarrow$ Su_S	11.5 (15)	17.9 (1)
	Su_B $\rightarrow$ Su_S_C	17.3 (15)	22.9 (1)
	SIFT_B $\rightarrow$ SIFT_S	3.1 (3)	11.7 (1)
	SIFT_B $\rightarrow$ SIFT_S_C	11.9 (15)	23.6 (1)
Fau.	M_B $\rightarrow$ M_S	27.7 (3)	32.4 (5)
	M_B $\rightarrow$ M_S_C	37.6 (15)	55.0 (1)
	H_B $\rightarrow$ H_S	29.2 (5)	52.5 (1)
	H_B $\rightarrow$ H_S_C	38.2 (10)	57.5 (1)
	KLT_B $\rightarrow$ KLT_S	2.4 (3)	24.0 (1)
	KLT_B $\rightarrow$ KLT_S_C	17.1 (5)	50.4 (1)
	Su_B $\rightarrow$ Su_S	25.7 (10)	40.6 (1)
	Su_B $\rightarrow$ Su_S_C	36.8 (5)	47.6 (1)
	SIFT_B $\rightarrow$ SIFT_S	<b>-0.8</b> (5)	23.2 (1)
	SIFT_B $\rightarrow$ SIFT_S_C	12.5 (10)	39.6 (1)

TAB. 3.2 – Gains moyens minimal et maximal sur l'erreur de validation pour les cinq images de référence et différents  $\Omega^{(f)}$  obtenus par les filtres (B), (S) et (C).

$I_{ref}$	$\Omega^{(i)}\_B \rightarrow \Omega^{(f)}$	g (%)   (a)	G (%)   (a)
Cam.	M_B $\rightarrow$ M_SFS	27.3 (10)	64.0 (1)
	M_B $\rightarrow$ M_SFS_C	32.7 (10)	62.4 (1)
	H_B $\rightarrow$ H_SFS	32.5 (10)	60.5 (1)
	H_B $\rightarrow$ H_SFS_C	34.2 (5)	55.9 (1)
	KLT_B $\rightarrow$ KLT_SFS	<b>20.1</b> (10)	47.8 (1)
	KLT_B $\rightarrow$ KLT_SFS_C	26.4 (5)	44.9 (1)
	Su_B $\rightarrow$ Su_SFS	29.7 (15)	76.0 (1)
	Su_B $\rightarrow$ Su_SFS_C	34.5 (15)	71.3 (1)
	SIFT_B $\rightarrow$ SIFT_SFS	20.5 (15)	53.4 (1)
	SIFT_B $\rightarrow$ SIFT_SFS_C	24.7 (15)	50.8 (1)
Art.	M_B $\rightarrow$ M_SFS	55.8 (15)	80.6 (1)
	M_B $\rightarrow$ M_SFS_C	55.5 (15)	78.9 (1)
	H_B $\rightarrow$ H_SFS	55.3 (15)	74.7 (1)
	H_B $\rightarrow$ H_SFS_C	52.3 (15)	73.0 (1)
	KLT_B $\rightarrow$ KLT_SFS	52.1 (15)	72.7 (1)
	KLT_B $\rightarrow$ KLT_SFS_C	49.3 (15)	73.1 (1)
	Su_B $\rightarrow$ Su_SFS	56.7 (15)	81.2 (1)
	Su_B $\rightarrow$ Su_SFS_C	57.9 (15)	81.2 (1)
	SIFT_B $\rightarrow$ SIFT_SFS	37.4 (15)	71.2 (1)
	SIFT_B $\rightarrow$ SIFT_SFS_C	39.6 (15)	71.8 (1)
Min.	M_B $\rightarrow$ M_SFS	-1016 (1)	44.4 (3)
	M_B $\rightarrow$ M_SFS_C	30.2 (15)	52.1 (1)
	H_B $\rightarrow$ H_SFS	37.4 (15)	65.2 (1)
	H_B $\rightarrow$ H_SFS_C	33.2 (15)	63.8 (1)
	KLT_B $\rightarrow$ KLT_SFS	25.0 (15)	56.1 (1)
	KLT_B $\rightarrow$ KLT_SFS_C	22.6 (15)	49.4 (1)
	Su_B $\rightarrow$ Su_SFS	27.7 (15)	66.1 (1)
	Su_B $\rightarrow$ Su_SFS_C	26.0 (15)	65.1 (1)
	SIFT_B $\rightarrow$ SIFT_SFS	28.3 (15)	63.7 (1)
	SIFT_B $\rightarrow$ SIFT_SFS_C	27.5 (15)	63.0 (1)
Noé	M_B $\rightarrow$ M_SFS	33.7 (15)	61.0 (1)
	M_B $\rightarrow$ M_SFS_C	32.3 (15)	58.7 (1)
	H_B $\rightarrow$ H_SFS	33.4 (15)	60.3 (1)
	H_B $\rightarrow$ H_SFS_C	33.8 (15)	62.8 (1)
	KLT_B $\rightarrow$ KLT_SFS	28.1 (15)	64.5 (1)
	KLT_B $\rightarrow$ KLT_SFS_C	28.8 (15)	57.2 (1)
	Su_B $\rightarrow$ Su_SFS	30.5 (15)	70.7 (1)
	Su_B $\rightarrow$ Su_SFS_C	30.0 (15)	65.0 (1)
	SIFT_B $\rightarrow$ SIFT_SFS	26.3 (15)	64.1 (1)
	SIFT_B $\rightarrow$ SIFT_SFS_C	28.2 (15)	65.8 (1)
Fau.	M_B $\rightarrow$ M_SFS	48.6 (15)	76.1 (1)
	M_B $\rightarrow$ M_SFS_C	45.6 (15)	76.3 (1)
	H_B $\rightarrow$ H_SFS	49.7 (15)	<b>84.6</b> (1)
	H_B $\rightarrow$ H_SFS_C	51.4 (15)	81.6 (1)
	KLT_B $\rightarrow$ KLT_SFS	39.1 (15)	76.9 (1)
	KLT_B $\rightarrow$ KLT_SFS_C	37.0 (15)	73.7 (1)
	Su_B $\rightarrow$ Su_SFS	47.9 (15)	79.5 (1)
	Su_B $\rightarrow$ Su_SFS_C	47.8 (15)	78.4 (1)
	SIFT_B $\rightarrow$ SIFT_SFS	29.4 (15)	65.4 (1)
	SIFT_B $\rightarrow$ SIFT_SFS_C	31.5 (15)	66.1 (1)

TAB. 3.3 – Gains moyens minimal et maximal sur l'erreur de validation pour les cinq images de référence et différents  $\Omega^{(f)}$  obtenus par les filtres (B), (SFS) et (C).

**B) Analyse**

- i) La figure 3.18-(a) (respectivement 3.18-(b)) montre, sur l'image « Noé » (respectivement « Cameraman ») et pour une amplitude  $a = 10$  pixels (respectivement  $a = 15$ ), l'intérêt d'opérer une étape de sélection ZV, que ce soit à partir des points de Moravec, de Harris et Stephens, de Shi et Tomasi (KLT), de Smith et Brady (SUSAN) ou de Lowe (SIFT). En effet, les lignes en pointillé sont toujours en-dessous des lignes correspondantes en points-tirets.

Plus précisément, pour « Noé » (figure 3.18-(a)), nous obtenons un gain de 29.3% sur M\_B, de 32.0% sur H\_B, de 28.0% sur KLT\_B, de 16.2% sur Su\_B et de 15.7% sur SIFT\_B. Alors que cette dernière valeur apparaît dans le tableau 3.1 (ligne 5, colonne 4) de « Noé », les quatre valeurs précédentes ne s'y trouvent pas car elles sont comprises entre les gains minimal et maximal.

Pour « Cameraman » (figure 3.18-(b)), nous obtenons un gain de 25.3% sur M\_B, de 27.9% sur H\_B (voir tableau 3.1 (ligne 2, colonne 4) de « Cameraman »), de 18.8% sur KLT\_B, de 17.5% sur Su\_B (voir tableau 3.1 (ligne 4, colonne 3) de « Cameraman ») et de 15.9% sur SIFT\_B.

Il en est souvent de même pour les autres images (voir tableau 3.1), sauf :

- à faible amplitude pour les images « Cameraman », « Artichaut » et « Minneapolis » et les points de Shi et Tomasi (KLT) et pour les images « Artichaut » et « Minneapolis » et les points de Lowe (SIFT) ;
- à forte amplitude pour l'image « Minneapolis » et les points de Moravec (voir les nombres en italique dans le tableau 3.1).

- 2i) La figure 3.19 illustre, sur « Artichaut » et pour  $a = 3$  pixels, l'intérêt d'opérer une étape S et encore mieux S\_C sur tout  $\Omega^{(i)}$  issu de Moravec (figure 3.19-(a)), de Harris et Stephens (figure 3.19-(b)), de Shi et Tomasi (KLT) (figure 3.19-(c)), de Smith et Brady (SUSAN) (figure 3.19-(d)) et de Lowe (SIFT) (figure 3.19-(e)).

En effet, les lignes en pointillé et en points-tirets sont toujours sous la ligne en trait plein de  $\Omega^{(i)}$ \_B. De plus, la ligne en pointillé correspondant à S\_C est plus souvent sous la ligne en points-tirets correspondant à S, excepté pour un ensemble final de 20 points.

Selon les images, les gains varient de -0.8% à 63.0% (voir tableau 3.2 et plus particulièrement les nombres en gras). Le gain de -0.8% est exceptionnel. C'est le seul gain moyen négatif. Il est obtenu pour l'image « Faugeras » et une amplitude de 5 pixels. Cette valeur exceptionnelle signifie juste qu'on est faiblement perdant dans ce cas-là de faire le filtre (S).

Par construction, le filtrage (C) conduit à une répartition uniforme à l'échelle des compartiments, ceci rend le prédicteur moins sensible aux problèmes d'occultation (voir les figures 3.15-(3-d), 3.15-(4-d) et 3.15-(5-d)).

- 3i) La figure 3.20 illustre, sur « Cameraman » et pour  $a = 10$  pixels, l'intérêt d'opérer une étape SFS et encore mieux SFS\_C sur tout  $\Omega^{(i)}$  issu de Moravec (figure 3.20-(a)), de Harris et Stephens (figure 3.20-(b)), de Shi et Tomasi (KLT) (figure 3.20-(c)), de Smith et Brady (SUSAN) (figure 3.20-(d)) et de Lowe (SIFT) (figure 3.20-(e)).

En effet, les lignes en pointillé et en points-tirets sont toujours sous la ligne en trait plein de  $\Omega^{(i)}$ \_B.

Selon les images, les gains varient de 20.1% à 84.6% (voir tableau 3.3 et plus particulièrement les nombres en gras). L'important gain négatif (en italique) dans ce tableau

est exceptionnel et a priori inexplicable. Il concerne le cas où  $\Omega^{(i)} = M$  et  $a = 1$  pixel pour l'image « Minneapolis ».

Selon les images, les gains de SFS peuvent être supérieurs ou très voisins des gains de SFS\_C. C'est le cas pour l'image « Artichaut » et  $\Omega^{(i)} = M$  ou H (voir tableau 3.3) contrairement à l'image « Cameraman » (voir figure 3.20).

- 4i) Les gains de SFS ou SFS\_C sont toujours bien supérieurs aux gains de tous les autres filtres (voir tableaux 3.1, 3.2 et 3.3).
- 5i) La figure 3.21 confirme l'absence de détecteur de points d'intérêt universel. Alors que pour « Faugeras », les points SIFT sont les plus pertinents, pour « Cameraman », ce sont les points SUSAN et pour « Artichaut », « Minneapolis » et « Noé », pas un seul ne s'impose.

### 3.7.3.2 Concernant le suivi de séquences artificielles

#### A) Des chiffres, des courbes

Le tableau 3.4 donne le nombre moyen d'images suivies  $i_{max}$  en fonction de :

- $\Omega^{(f)}$  associé à 60 points d'intérêt,
- $a = 15$  pixels,
- $I_{ref}$ ,
- la hiérarchie des suivis enchaînés.

Pour chaque image, les  $\Omega^{(f)}$  sont classés, de haut en bas, dans l'ordre des erreurs de validation (EV) croissantes.

Image	Sélection $\Omega^{(f)}$	EV	S15	S15-10 -5-3-1
« Faugeras »	H_SFS_C	4.5	27.55	40.90
	Su_S_C	5.3	22.90	37.30
	SIFT_ZV_S_C	5.8	30.85	53.50
	H_B	9.5	10.25	14.75
« Cameraman »	H_SFS_C	4.8	54.10	97.65
	H_ZV_S_C	5.5	47.10	81.20
	M_S_C	6.0	44.80	81.25
	M_B	8.7	15.55	40.85
« Minneapolis »	H_SFS	4.1	47.45	80.65
	H_S_C	5.2	50.55	84.50
	H_ZV_S_C	5.5	36.20	71.65
	M_ZV	7.3	25.35	49.00
« Noé »	M_SFS	5.2	36.50	67.65
	KLT_ZV_S_C	6.5	35.30	77.60
	SIFT_S	7.0	31.25	67.80
	H_B	8.3	14.20	32.30
« Artichaut »	M_SFS_C	2.4	39.80	62.65
	M_S_C	3.5	40.85	76.85
	M_ZV_S_C	3.6	38.80	70.15
	Su_B	7.0	13.40	43.85

TAB. 3.4 – Nombre moyen d'images suivies  $i_{max}$  sur 20 séquences pour chaque type d'images et différents  $\Omega^{(f)}$  ( $m = 60$ ).

**B) Analyse**

i) Meilleur est  $\Omega^{(f)}$  (au sens de l'erreur de validation), plus résistant est, en général, l'algorithme de suivi. Ceci est illustré dans la colonne 4 du tableau 3.4. Ce commentaire n'est que partiellement vrai pour les séquences :

- « Faugeras » avec  $\Omega^{(f)} = \text{Su\_S\_C}$  et  $\Omega^{(f)} = \text{H\_SFS\_C}$ ,
- « Minneapolis » avec  $\Omega^{(f)} = \text{H\_SFS}$ ,
- « Artichaut » avec  $\Omega^{(f)} = \text{M\_SFS\_C}$ .

Ceci montre certainement l'inadéquation du prédicteur  $\hat{A}$  !

2i) La colonne 5 du tableau 3.4 montre, quant à elle, l'utilité du suivi hiérarchique. En effet, le nombre moyen d'images suivies est toujours plus grand dans des proportions variant de 1.4 (pour « Faugeras » et  $\Omega^{(f)} = \text{H\_B}$ ) à 3.3 (pour « Artichaut » et  $\Omega^{(f)} = \text{Su\_B}$ ).

3i) Les figures 3.16 et 3.17 et le tableau 3.4 montrent que la sélection d'un bon ensemble de points est nécessaire. En effet, pour un suivi avec une seule hiérarchie, un bon ensemble de points permet de suivre, en moyenne, le nombre d'images pour lesquelles les amplitudes maximales à suivre sont de l'ordre de grandeur des amplitudes apprises. Ainsi, pour la séquence d'images artificielle présentée sur la figure 3.16 et pour une amplitude de 15 pixels :

- à partir de l'image 49, les amplitudes maximales à suivre sont supérieures à 15 pixels,
- le nombre d'images suivies est de 21 pour H\_B et de 43 pour M\_SFS.

Le suivi hiérarchique permet de suivre au-delà de l'image 49.

**3.8 Bilan**

Dans ce chapitre, notre intérêt s'est porté sur la détermination du meilleur sous-ensemble de points d'intérêt afin d'améliorer l'algorithme de suivi d'objets plans indéformables de Jurie et Dhome.

Il est à noter que plus il y a de points d'intérêt de qualité, meilleur c'est.

Après avoir établi un historique des détecteurs de points d'intérêt de 1976 à nos jours, nous avons émis l'idée d'utiliser des approches de statistique et d'algèbre linéaire pour sélectionner des sous-ensembles de ces points et non des points indépendamment les uns des autres.

Notons que notre démarche, bien qu'assez systématique, n'est naturellement pas exhaustive. Des détecteurs récents comme SURF (en anglais, *Speeded Up Robust Features*) par exemple pourraient rejoindre notre liste déjà importante.

Nous avons montré l'utilité de sélectionner des points d'intérêt en adéquation avec l'algorithme de suivi et la dépendance entre le meilleur détecteur de points d'intérêt et la texture de l'image traitée.

Enfin, nous recommandons de ne pas prendre les meilleurs points au sens du critère de détection mais de faire, quoiqu'il arrive, une sélection par compartiments et d'appliquer le filtre (SFS) concernant l'approche statistique de sélection séquentielle par ajout qui donne les meilleurs résultats. L'inconvénient de ce filtre est sa complexité algorithmique bien supérieure aux autres. Le second filtre à privilégier est le filtre (S) concernant l'approche algébrique de sélection d'un sous-ensemble.





## Chapitre 4

# Suivi d'objets déformables

L'outil mathématique **maillage**, présenté et utilisé dans ce chapitre, est fondamental en traitement d'images et en vision. Un maillage peut être vu comme une représentation géométrique discrète d'un domaine continu fort utile pour :

- représenter des **objets déformables** (définition 2.1.1.2) ou des régions denses de manière compacte, c'est-à-dire avec peu de données, associées aux nœuds du maillage et nommées **valeurs nodales** ;
- reconstruire une approximation continue de la réalité via une **interpolation** ;
- faire des analyses multi-résolutions et multi-échelles des objets, du mouvement...
- faire des analyses spatio-temporelles paramétriques des déformations.

En vision, les maillages sont couramment utilisés sans faire forcément référence à leur propriété (conformité, topologie, géométrie, compacité...), sans étudier de manière exacte l'inversion et la composition de leurs transformations. Nous nous attelons dans ce chapitre à présenter tout cela dans le cadre du **suivi d'objets déformables**.

Nos contributions sont :

- 1) la caractérisation analytique d'une transformation nodale ;
- 2) la caractérisation géométrique de l'inversion exacte d'une transformation nodale ;
- 3) l'approximation d'une composition d'une transformation globale et d'une transformation nodale en utilisant les **coordonnées barycentriques généralisées** ;
- 4) la proposition d'un nouvel algorithme de suivi fondé sur l'apprentissage d'un prédicteur linéaire par nœud du maillage.

Ce chapitre s'articule comme suit :

- des généralités sur les maillages ;
- un bref état de l'art sur les méthodes de suivi par maillage ;
- notre algorithme de suivi d'objets déformables ;
- une étape d'expérimentation ;
- un bilan.

### 4.1 Généralités sur les maillages

#### 4.1.1 Définitions relatives aux maillages

Les principales définitions présentées dans cette section proviennent des ouvrages de Paul-Louis George [Geo91, GB97] et du mémoire de thèse de Gwénaëlle Marquant [Mar00].

#### 4.1.1.1 Définition (Maillage)

Un **maillage**  $\mathcal{M}$  est une division d'un domaine  $\Omega \subset \mathbb{R}^2$  en polygones  $F_i$ , appelés **faces** ou **mailles**, telle qu'en notant  $\overset{\circ}{F}_i = F_i \setminus \partial F_i$  l'intérieur de la face  $F_i$  :

$$\begin{cases} \Omega = \bigcup_{i=1}^K F_i \\ \forall i, \overset{\circ}{F}_i \neq \emptyset \\ \forall i \neq j, \overset{\circ}{F}_i \cap \overset{\circ}{F}_j = \emptyset \end{cases} \quad (4.1)$$

L'ensemble  $\mathcal{S}$  des sommets  $(S_j)_{j=1..N}$  des faces forme l'ensemble des sommets, aussi appelés **nœuds**, du maillage.

L'ensemble  $\mathcal{A}$  des frontières  $(A_k)_{k=1..L}$  des faces forme l'ensemble des **arêtes** du maillage.

Tandis que la figure 4.1 présente un exemple de maillage, la figure 4.2 illustre, quant à elle, les différents composants (faces, arêtes, nœuds) d'un maillage.

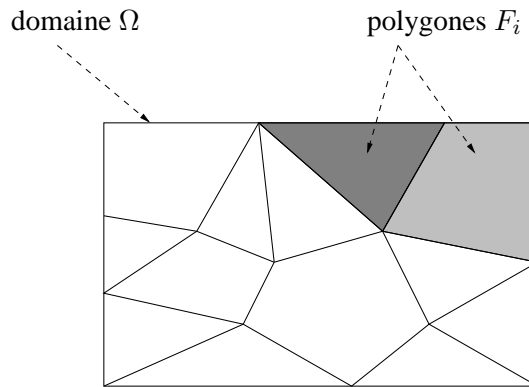


FIG. 4.1 – Exemple de maillage.

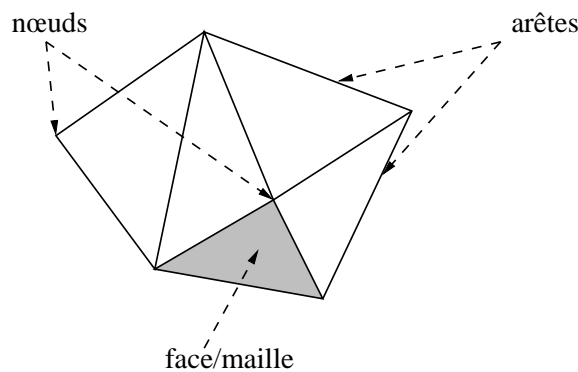


FIG. 4.2 – Les nœuds, les arêtes et les faces définissent un maillage.

Même si, en théorie, les polygones de la division sont quelconques, en pratique, dans un maillage, tous les polygones sont du même ordre. Les **maillages triangulaire, quadrangulaire** et **hexagonal** ne sont formés respectivement que de polygones à trois, quatre et six

côtés (c'est-à-dire d'ordres trois, quatre et six, donc des triangles, des quadrilatères et des hexagones) (voir figure 4.3). Les maillages les plus utilisés sont les maillages triangulaires et quadrangulaires. Dans ce mémoire de thèse, nous ne considérons que des maillages triangulaires. Ils sont obtenus par une triangulation de Delaunay que nous expliquons à la section 4.1.2.2.

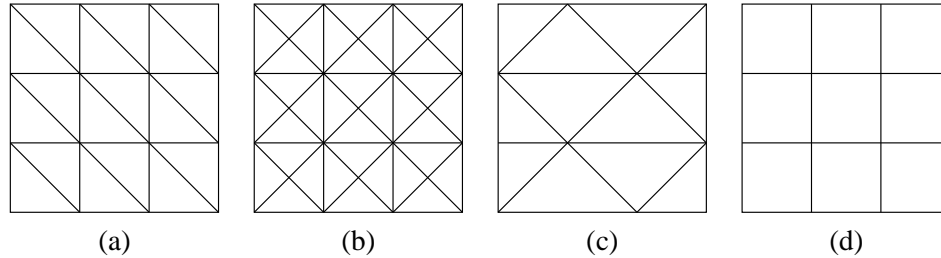


FIG. 4.3 – Exemples de maillages triangulaires (a), (b) et (c) et quadrangulaire (d).

#### 4.1.1.2 Définition (Géométrie d'un maillage)

*La position des nœuds d'un maillage définit sa **géométrie**.*

#### 4.1.1.3 Définition (Topologie d'un maillage)

*Les arêtes d'un maillage définissent sa **topologie**.*

À titre d'exemple, les maillages (a) et (d) de la figure 4.3 possèdent la même géométrie, mais des topologies différentes.

#### 4.1.1.4 Définition (Maillage conforme)

*Un maillage est dit **conforme** si l'intersection de deux arêtes distinctes est :*

- un nœud,
- ou l'ensemble vide,

*c'est-à-dire :*

$$\forall i \neq j, A_i \cap A_j \in \{\{S_n\}_{n=1..N}\} \cup \{\emptyset\}. \quad (4.2)$$

*sinon il est dit **non conforme**.*

Il est à noter qu'à la construction le choix est toujours de générer un maillage conforme. Il pourra devenir non conforme lors du déplacement d'un de ses nœuds. Cette propriété de conformité est une condition suffisante assurant la continuité des polynômes d'interpolation (voir §4.1.3) d'une face à l'autre. Tous les maillages de la figure 4.4 sont conformes dans le cadre d'un maillage composé de polygones d'ordre quelconque. Par contre, seul le maillage 4.4-(b) est non conforme dans le cadre d'un maillage triangulaire.

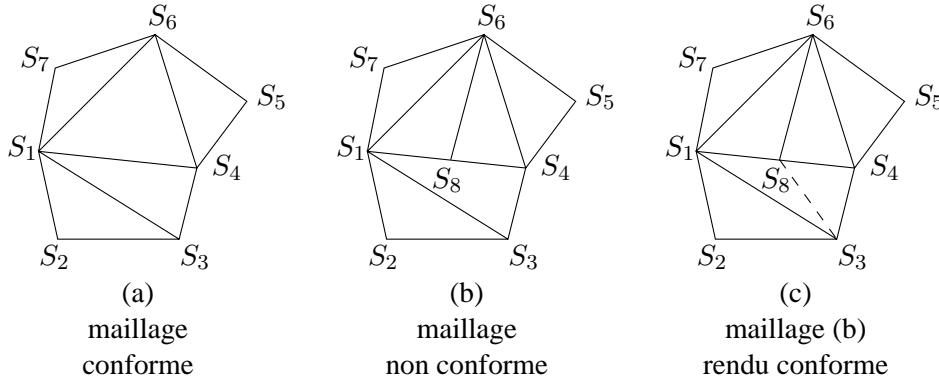


FIG. 4.4 – Exemples de maillages conformes (a) et (c) et non conforme (b). En effet,  $[S_1S_8] \cap [S_1S_4] = [S_1S_8] \notin \{\{S_1\}, \{S_2\}, \{S_3\}, \{S_4\}, \{S_5\}, \{S_6\}, \{S_7\}, \{S_8\}\} \cup \{\emptyset\}$ , par exemple. Le maillage (b) est rendu conforme par l'ajout de l'arête en tiret du maillage (c).

#### 4.1.1.5 Définition (Maillage régulier/irrégulier)

Un maillage composé de polygones identiques est dit **régulier** sinon il est dit **irrégulier**.

Tandis que les figures 4.3-(a), (b) et (d) illustrent un maillage régulier, les maillages représentés sur les figures 4.3-(c), 4.5 et 4.6 sont irréguliers.

En traitement numérique des images, un maillage irrégulier est bien adapté dans certaines situations comme l'analyse multi-résolution de la texture, par exemple. Il peut être obtenu de deux façons au moins :

- en sélectionnant les nœuds du maillage de manière plus dense sur le contour des objets (voir figure 4.5),

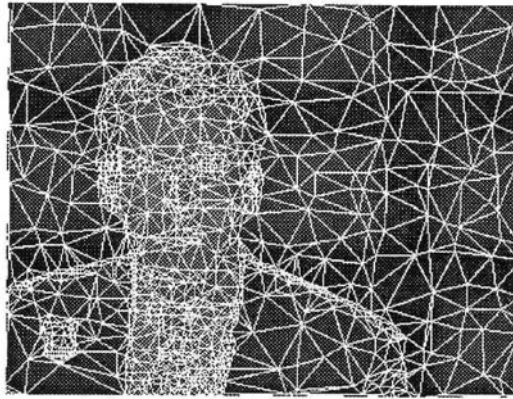


FIG. 4.5 – Exemple de maillage irrégulier obtenu à partir d'un nuage de points plus dense sur les contours des objets (image du mémoire de thèse de Pierre Alliez [[All00], fig. 3.50, p. 116]).

- en déformant un maillage régulier. Par exemple, en déplaçant de manière « optimale » les nœuds, au cours d'un traitement de décimation (voir §4.1.2.3) ou de subdivision (voir §4.1.2.3) adapté au contenu de l'image (voir figure 4.6).

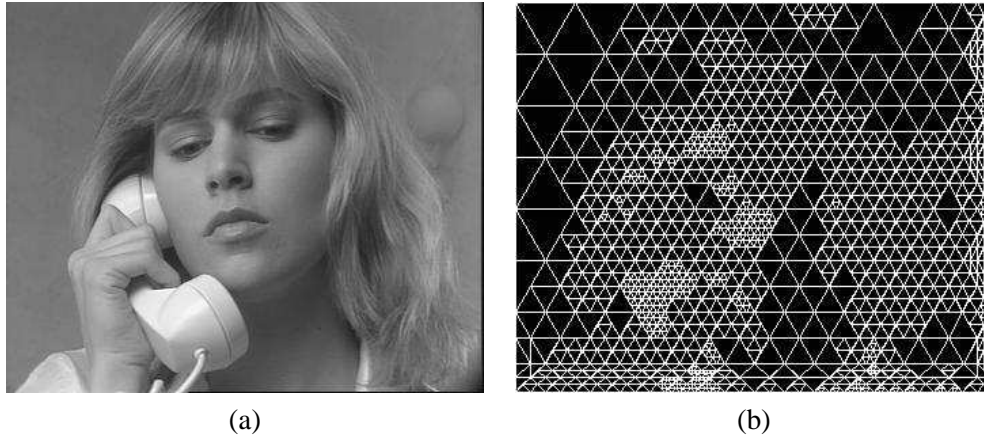


FIG. 4.6 – (a) Image initiale (Susie) ; (b) maillage non conforme adapté à l'image (images du mémoire de thèse de Gwenaëlle Marquant [[Mar00], fig. 6.8, p. 139]).

#### 4.1.1.6 Définition (Compacité)

La **compacité**  $C$  d'un polygone est donnée par la formule :

$$C = \frac{4\pi S}{P^2}, \quad (4.3)$$

où  $S$  est la surface du polygone et  $P$  son périmètre.

La compacité d'un polygone est maximale quand il est régulier (c'est-à-dire un polygone à côtés et angles égaux). Dans le cas des triangles, la compacité est donc maximale pour un triangle équilatéral. La figure 4.7 montre la valeur de la compacité d'un triangle  $\overline{S_1 S_2 S_3}$  en fonction de la position du sommet  $S_3$  sur l'axe des abscisses. Les sommets  $S_1$ ,  $S_2$  et  $S_3$  ont pour coordonnées respectives  $[0\ 0]^T$ ,  $[50\ 50\sqrt{3}]^T$  et  $[x\ 0]^T$ . La compacité est maximale pour  $x = 100$ , c'est-à-dire lorsque le triangle est équilatéral. Cette mesure est utilisée en reconnaissance de formes. Elle permet de caractériser en partie la forme d'un objet.

#### 4.1.1.7 Définition (Maillage statique/dynamique)

Un maillage est dit **statique** si la position de ses nœuds est invariante dans le temps, sinon il est dit **dynamique**.

Un exemple d'utilisation des maillages statiques est la représentation compressée d'une image (voir figures 4.5 et 4.6). Un exemple d'utilisation des maillages dynamiques est la représentation synthétique d'un champ de mouvement entre deux images. Dans ce mémoire de thèse, nous utilisons uniquement les maillages dynamiques.

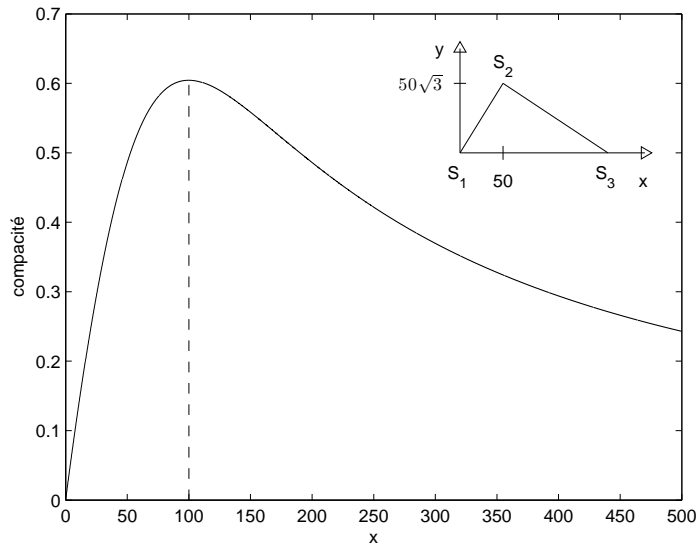


FIG. 4.7 – Compacité d'une maille triangulaire de sommets  $S_1$ ,  $S_2$  et  $S_3$  de coordonnées respectives  $[0\ 0]^T$ ,  $[50\ 50\sqrt{3}]^T$  et  $[x\ 0]^T$  en fonction de la position  $x$  du sommet  $S_3$  sur l'axe des abscisses.

#### 4.1.1.8 Définition (Polygone support d'un nœud)

Soient :

- $\mathcal{M}$  un maillage,
- $S_n$  un nœud de  $\mathcal{M}$ ,
- $\{F_k\}_{k=1..K}$  les faces de  $\mathcal{M}$ ,

alors le **polygone support**  $\Pi_n$  de  $S_n$  est défini par :

$$\Pi_n = \bigcup_{k \in [1, K]_{\mathbb{N}} | S_n \in F_k} F_k \quad (4.4)$$

Une illustration du polygone support est donnée dans la figure 4.8. L'intérêt du polygone support en traitement numérique des images est de définir un voisinage (au sens de contexte) du nœud  $S$  qui varie avec ses déplacements.

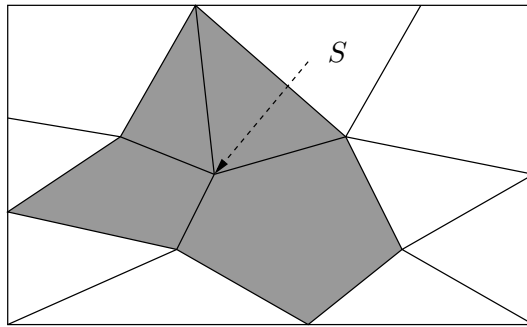


FIG. 4.8 – En gris, le polygone support du nœud  $S$ .

#### 4.1.1.9 Définition (Retournement de maille)

On parle de **retournement de maille** lorsqu'un nœud du maillage est déplacé en dehors de la frontière de son polygone support.

La figure 4.9 illustre ce phénomène. Un retournement de maille conduit toujours à un maillage non conforme.

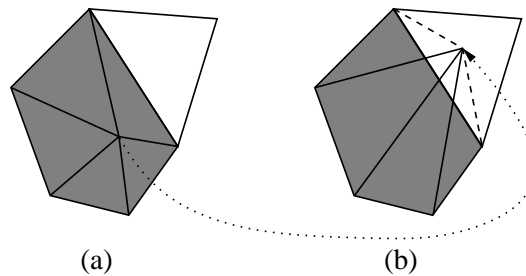


FIG. 4.9 – Exemple de retournement de maille : (a) le maillage initial ; la flèche en pointillé indique le déplacement du nœud ; (b) la maille résultante (en tiret) est retournée.

#### 4.1.1.10 Définition (Valeur nodale)

Une **valeur nodale** est un attribut numérique (scalaire ou vectoriel) associé à un nœud du maillage.

Cette valeur dépend de l'utilisation du maillage. Dans le cas du codage d'une image, la valeur nodale est un niveau de gris. Pour le suivi, c'est un vecteur de déplacement. Enfin, dans le cas d'une représentation 3D, elle peut être le vecteur normal à la surface de la maille, sa couleur... Un **polynôme d'interpolation** (voir §4.1.3) permet d'attribuer des valeurs en tout point d'une face du maillage à partir des valeurs nodales de ses sommets.

### 4.1.2 Topologie et géométrie d'un maillage

Cette section est consacrée à la construction et la modification d'un maillage. Nous développons un « petit » état de l'art sur la **sélection des nœuds** :

- la célèbre **partition de Voronoï**,
- la non moins célèbre **triangulation de Delaunay**.

Nous décrivons quelques opérateurs de modification de la géométrie et de la topologie d'un maillage : la **décimation** et la **subdivision**.

#### 4.1.2.1 Choix de la géométrie d'un maillage

La géométrie d'un maillage, c'est-à-dire la position de ses nœuds, est importante. Elle peut être :

- **indépendante de l'image** : les nœuds sont uniformément répartis sur toute l'image ;
- **dépendante de l'image** : de nombreuses propositions ont été faites et nous en présentons quelques-unes.

### A) Algorithme de Niewęglowski *et al.*

Cet algorithme [NCH93] répartit les nœuds au voisinage des contours des objets en deux temps :

- i) sélection des nœuds indépendante de l'image,
- 2i) pour chaque nœud  $S$  issu de (i), remplacement par le voisin qui a la meilleure réponse au détecteur de contour choisi.

La géométrie « indépendante » initiale est donc déformée afin que chacun de ses nœuds soit déplacé sur son contour le plus proche.

### B) Algorithme de Wang et Lee

Cet algorithme [[WL94], §III, p. 613] trouve la position de tous les nœuds  $S_n$ ,  $n \in [1, N]_{\mathbb{N}}$  qui sont solution du problème d'optimisation  $(\mathcal{P}^{wl})$ .  $(\mathcal{P}^{wl})$  est résolu à l'aide de la méthode de la plus forte pente (*steepest descent*) [[NW06], p. 21] :

$$(\mathcal{P}^{wl}) \quad \begin{cases} \min \\ \mathcal{S} = \{S_n\}_{n=1..N} \end{cases} E(\mathcal{S}) = w_i E_i(\mathcal{S}) + w_p E_p(\mathcal{S}) + w_d E_d(\mathcal{S}) \quad , \quad (4.5)$$

où :

- $E_i$ , énergie d'interpolation globale [[WL94], éq. (12), p. 614] vaut :

$$E_i(\mathcal{S}) = \frac{1}{2} \sum_{(x,y) \in \Omega} (I(x,y) - \hat{I}(x,y))^2, \quad (4.6)$$

où  $\Omega$  est le domaine du maillage défini par  $\mathcal{S}$ ,  $I(x,y)$  est l'intensité de l'image  $I$  en  $(x,y)$  et  $\hat{I}(x,y)$  est l'intensité en  $(x,y)$  interpolée à partir des valeurs nodales, ici des niveaux de gris, associées aux sommets de la maille contenant  $(x,y)$ .  $E_i$  est homogène au carré d'un niveau de gris ;

- $E_p$ , énergie de position d'intérêt, contraint les nœuds à être sur les contours et les coins [[WL94], éq. (10), p. 614].  $E_p$  est sans unité ;
- $E_d$ , énergie de déformation [[WL94], éq. (7), p. 613], contraint le maillage en terme de régularité, c'est-à-dire à avoir une compacité maximale (définition 4.1.1.6).  $E_d$  est homogène au carré d'une distance exprimée en pixel ;
- $w_i$ ,  $w_p$  et  $w_d$  pondèrent les énergies associées pour obtenir trois termes aux mêmes ordres de grandeur. Naturellement,  $w_i + w_p + w_d = 1$ .

### C) Algorithme de Altunbasak et Tekalp

Dans le cadre du traitement de l'image  $I_t$  d'une séquence  $\{I_0, \dots, I_{t-1}, I_t, \dots\}$ , cet algorithme sélectionne des nœuds qui :

- i) contraignent les frontières du maillage à s'aligner sur celles de(s) l'objet(s) d'intérêt,
- 2i) sont plus nombreux sur les contours et particulièrement dans les zones de l'image  $I_t$  déformées par rapport à  $I_{t-1}$ .

L'algorithme 4.1.1 reprend celui de Altunbasak et Tekalp [[AT97a], §D, p. 1273 et fig. 4, p. 1274].



- 1)  $\mathcal{S} \leftarrow \emptyset$  - initialement, l'ensemble  $\mathcal{S}$  des nœuds du maillage est vide
  - 2) Calculer le champ dense  $\mathcal{CM}$  de mouvement entre deux images  $I_{t-1}$  et  $I_t$  en appliquant l'algorithme de Lucas-Kanade [LK81] ou de Horn-Schunck [Hor86].
  - 3) Labéliser NM (« Non Marqué ») tous les pixels appartenant à l'objet d'intérêt : on obtient  $\Omega_{NM}$ .
  - 4)
$$\forall p = (x, y) \in \Omega_{NM}, C(p) \leftarrow \left| \frac{\partial I}{\partial x}(p) \right| + \left| \frac{\partial I}{\partial y}(p) \right|. \quad (4.7)$$
  - 5) **Tant que**  $|\Omega_{NM}| \neq 0$  **Faire**
    - 5.1)  $\mathcal{S} \leftarrow \mathcal{S} \cup \{P\}$ , où :
$$P = \left\{ \underset{p \in \Omega_{NM}}{\operatorname{argmax}} \quad C(p) \right\}. \quad (4.8)$$
    - 5.2)
$$E_m \leftarrow \frac{1}{|\Omega_{NM}|} \sum_{p \in \Omega_{NM}} (I_t(p) - \hat{I}_t(p))^2, \quad (4.9)$$

où  $\hat{I}_t$  est l'image prédite par  $\mathcal{CM}$ .
    - 5.3) Déterminer le disque  $\mathcal{D}$  de centre  $P$  tel que :
$$\sum_{p \in \mathcal{D}} (I_t(p) - \hat{I}_t(p))^2 \approx E_m. \quad (4.10)$$
    - 5.4) Labéliser M (« Marqué ») tous les pixels de  $\mathcal{D}$ , d'où :
$$\Omega_{NM} \leftarrow \Omega_{NM} \setminus \mathcal{D}.$$
- Fin Tant que**

Algorithme 4.1.1: Algorithme de sélection des nœuds d'un maillage proposé par Altunbasak et Tekalp [[AT97a], §D, p. 1273].

#### D) Algorithme de Raphaëlle Balter

Raphaëlle Balter [[Bal05], §1.2.1, p. 134] compartimente l'image en blocs et applique sur chaque bloc le détecteur de points d'intérêt de Harris (voir §3.2.4). Elle obtient ainsi une répartition plus uniforme des nœuds du maillage.

##### 4.1.2.2 Construction d'un maillage plan 2D

À partir d'un ensemble de points  $\mathcal{S}$  d'un domaine plan  $\Omega$ , le diagramme de Voronoï et la triangulation de Delaunay construisent tous deux un maillage plan conforme. Alors que dans le diagramme de Voronoï, les mailles sont quelconques, dans la triangulation de Delaunay, ce sont des triangles. Dans ce mémoire de thèse, les maillages triangulaires découlent d'une triangulation de Delaunay.

### A) Diagramme de Voronoï

Le diagramme de Voronoï doit son nom au mathématicien Georgy Fedoseevich Voronoï (1868-1908).

#### 4.1.2.2.1 Définition (Diagramme de Voronoï)

Soit  $\mathcal{S}$  un ensemble de points d'un domaine plan  $\Omega$ , alors le **diagramme de Voronoï** de  $\Omega$ , défini à partir de  $\mathcal{S}$ , est un ensemble de régions convexes appelées **régions de Voronoï**.

#### 4.1.2.2.2 Définition (Région de Voronoï)

Soit  $\mathcal{S} = \{S_i\}_{i=1..N}$  un ensemble de points d'un domaine plan  $\Omega$ , la **région de Voronoï**  $V_i$  associée au point  $S_i$  est définie comme suit :

$$V_i = \{X \in \Omega | \forall j \neq i, \|X - S_i\|_2 \leq \|X - S_j\|_2\}. \quad (4.11)$$

La figure 4.10-(b) montre la partition de Voronoï associée à l'ensemble de points illustré sur la figure 4.10-(a).

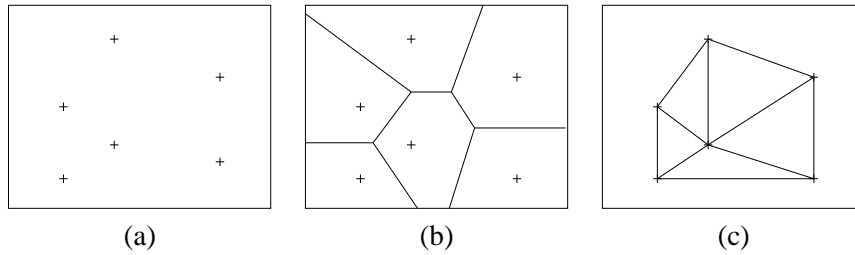


FIG. 4.10 – (a) Nuage de points  $\mathcal{S}$  ; (b) partition de Voronoï ; (c) triangulation de Delaunay.

### B) Triangulation de Delaunay

La triangulation de Delaunay est une méthode qui a été développée dans les années 1930. Elle doit son nom au mathématicien Boris Nikolaevich Delone (1890-1980) et prolonge les résultats de Voronoï. Elle génère un maillage triangulaire. Elle aboutit à un graphe dit graphe de Delaunay, souvent qualifié de « dual » du diagramme de Voronoï au sens où : une arête relie deux sommets  $S_1$  et  $S_2$  si et seulement si l'intersection des régions de Voronoï  $V_1$  et  $V_2$  est une arête. La figure 4.10-(c) donne la triangulation de Delaunay associée au nuage de points de la figure 4.10-(a).

#### 4.1.2.2.3 Définition (Triangulation de Delaunay)

Soit  $\mathcal{S} = \{S_i\}_{i=1..n}$  un ensemble de points d'un domaine plan  $\Omega$ .

Une triangulation  $\mathcal{D}$  est une **triangulation de Delaunay** de  $\mathcal{S}$  si pour tout triangle  $(t)$  de  $\mathcal{D}$ , son cercle circonscrit  $\mathcal{C}$  ne contient aucun point de  $\mathcal{S}$  en son intérieur, ce qui équivaut en termes mathématiques à :

$\forall t \in \mathcal{D}$ , en notant  $D$  le disque tel que  $\partial D = \mathcal{C}$ , alors  $\forall S \in \mathcal{S}, S \notin D \setminus \mathcal{C}$ .

Parmi les propriétés de la triangulation de Delaunay [[GB97], §3.2, p. 31, p. 49, p. 51], citons :

- la triangulation de Delaunay est unique si, parmi tous les sommets considérés, il n'y a pas quatre sommets cocycliques ;
- parmi les triangulations possibles du nuage de points, la triangulation de Delaunay maximise l'angle minimal entre deux arêtes et minimise le rayon maximal des cercles circonscrits aux triangles ;
- les triangles générés par la triangulation de Delaunay ont une compacité (définition 4.1.1.6) maximale.

### C) Triangulation de Delaunay avec contraintes

La triangulation de Delaunay avec contraintes est construite à partir d'un ensemble de sommets et d'un ensemble d'arêtes qui ne se coupent qu'en des sommets. Ces arêtes sont les contraintes et sont partie intégrante de la triangulation. La triangulation de Delaunay avec contraintes vérifie au mieux les propriétés de la triangulation de Delaunay. Un exemple d'utilisation de cette triangulation est celle d'un objet où les arêtes de son contour forment l'ensemble des contraintes.

#### 4.1.2.2.4 Définition (Triangulation de Delaunay avec contraintes)

Soit  $\mathcal{S} = \{S_i\}_{i=1..n}$  un ensemble de points d'un domaine plan  $\Omega$ . Soit  $\mathcal{A} = \{A_i\}_{i=1..m}$  un ensemble (éventuellement vide) d'arêtes d'extrémités des points de  $S$  tels que :

$$\forall i \neq j, a_i \cap a_j \in \{\{S\}_{S \in \mathcal{S}}\} \cup \{\emptyset\}. \quad (4.12)$$

Une triangulation  $\mathcal{D}_C$  est une **triangulation de Delaunay avec contraintes** de  $\mathcal{S}$  si :

- les arêtes contraintes appartiennent aux arêtes de la triangulation,
- pour tout triangle  $(t)$  de  $\mathcal{D}_C$ , si son cercle circonscrit  $\mathcal{C}$  contient un point  $S$  quelconque de  $\mathcal{S}$  en son intérieur (c'est-à-dire si la caractérisation de la triangulation de Delaunay n'est pas vérifiée), alors  $S$  est caché des sommets de  $t$ , c'est-à-dire que n'importe quel segment d'extrémités  $S$  et un sommet de  $t$  intersecte une arête contrainte,

ce qui équivaut en termes mathématiques à :

- $\forall a \in \mathcal{A}, a \in \mathcal{D}_C$
- $\forall t \in \mathcal{D}_C, (t, \text{triangle de sommets } S_i, S_j \text{ et } S_k), \text{ en notant } D \text{ le disque tel que } \partial D = \mathcal{C}, \text{ alors :}$

$$\forall S \in \mathcal{S}, S \in D \setminus \mathcal{C} \Rightarrow \begin{cases} \exists a_l \in \mathcal{A} \text{ tel que } [SS_i] \cap a_l \neq \emptyset \\ \exists a_l \in \mathcal{A} \text{ tel que } [SS_j] \cap a_l \neq \emptyset \\ \exists a_l \in \mathcal{A} \text{ tel que } [SS_k] \cap a_l \neq \emptyset \end{cases}$$

### D) Algorithmes de triangulation de Delaunay

La triangulation de Delaunay peut être, entre autres, obtenue :

- à partir de la partition de Voronoï ;
- via une méthode itérative [[dBvKOS00], algorithme, p. 192], [[GB97], §2.4, p. 55] qui insère un nouveau sommet dans la triangulation à chaque itération ;
- via une méthode récursive [[Che87], p. 217–222], [[GB97], §2.5.2, p. 61] détaillée dans l'algorithme 4.1.2.

```

fonction  $\mathcal{D} = \text{triangulation}(\mathcal{S})$ 
   $n \leftarrow |\mathcal{S}|$ 
  cas  $n=1$  :
     $\mathcal{D} \leftarrow S$ , où  $\mathcal{S} = \{S\}$ 
  cas  $n=2$  :
     $\mathcal{D} \leftarrow [S_1 S_2]$ , où  $\mathcal{S} = \{S_1, S_2\}$ 
  cas  $n=3$  :
     $\mathcal{D} \leftarrow \widehat{S_1 S_2 S_3}$ , où  $\mathcal{S} = \{S_1, S_2, S_3\}$ 
  autres cas :
    Choisir arbitrairement  $\mathcal{S}_1$  et  $\mathcal{S}_2$  tels que :
       $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ ,  $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$ ,  $\mathcal{S}_1 \neq \emptyset$  et  $\mathcal{S}_2 \neq \emptyset$ 
     $\mathcal{D}_1 \leftarrow \text{triangulation}(\mathcal{S}_1)$ 
     $\mathcal{D}_2 \leftarrow \text{triangulation}(\mathcal{S}_2)$ 
     $\mathcal{D} \leftarrow \text{fusion}(\mathcal{D}_1, \mathcal{D}_2)$ 
  Fin cas

```

Algorithme 4.1.2: Algorithme récursif de construction d'une triangulation de Delaunay.

Pour les deux derniers algorithmes, la triangulation est effectuée en un temps optimal de  $O(n \log(n))$ , où  $n$  est le nombre de sommets.

Le logiciel libre « Triangle » [She] créé par Shewchuk à l'université de Carnegie Mellon permet de générer, entre autres, des triangulations de Delaunay avec ou sans contraintes en utilisant les algorithmes de type « itératif » et « récursif ».

## E) Synthèse

En raison des bonnes propriétés de la triangulation de Delaunay et de sa faible complexité algorithmique, nous l'avons choisie dans tous nos tests.

### 4.1.2.3 Modifications de la géométrie ou de la topologie d'un maillage

On modifie la géométrie ou la topologie d'un maillage (par exemple, pour l'adapter au contenu de l'image) :

- en déplaçant des nœuds,
- en supprimant des nœuds (des arêtes, des faces) : c'est la **décimation**,
- en ajoutant des nœuds (des arêtes, des faces) : c'est la **subdivision**.

#### A) La décimation

Appliquée itérativement, cette opération permet d'atteindre un objectif en termes de complexité géométrique.

Six opérateurs élémentaires de simplification permettent de réaliser une décimation (voir [[All00], §2.2.1, p. 29]) :

- **suppression d'un nœud** : la figure 4.11 illustre cette opération qui nécessite, dans le cas d'un maillage triangulaire, une phase de re-triangulation. Il est à noter que la nouvelle triangulation n'est pas nécessairement unique.

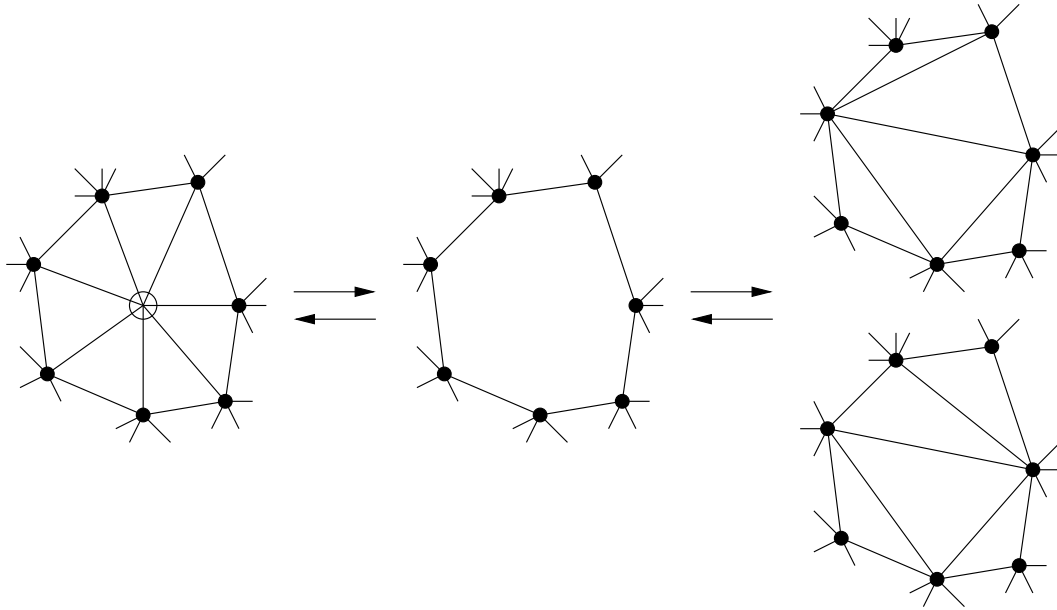
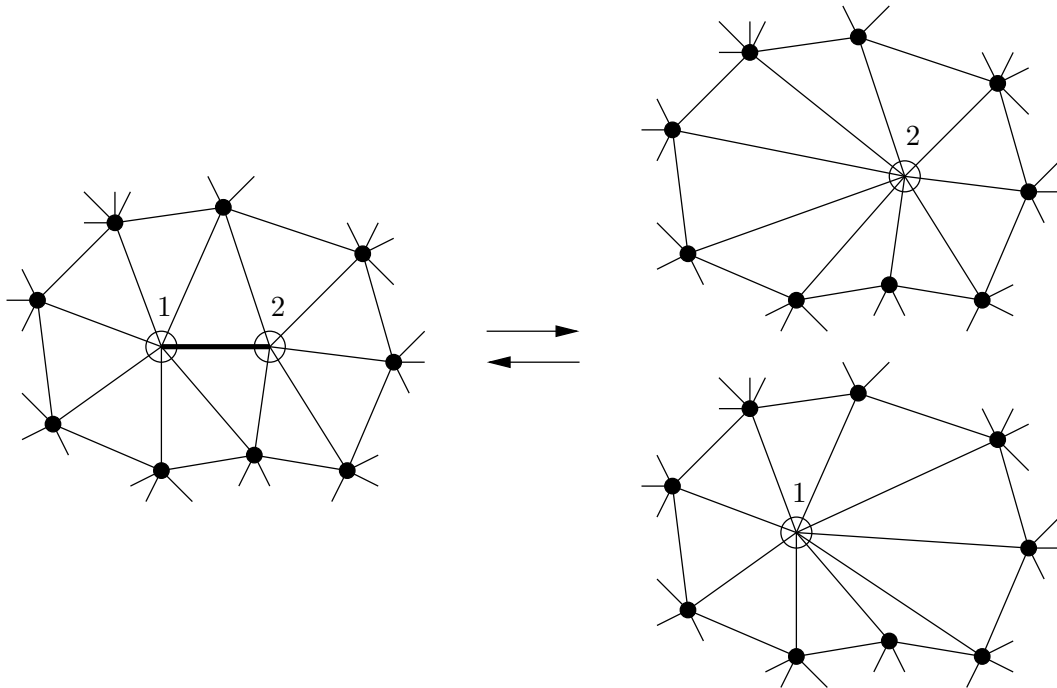
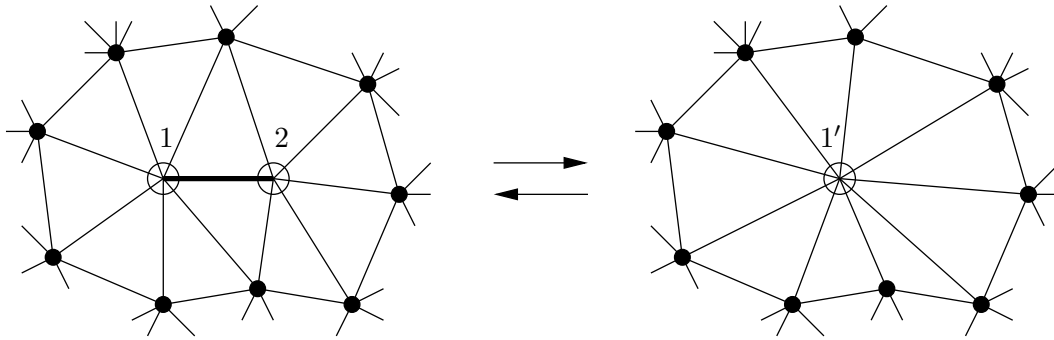


FIG. 4.11 – Suppression d'un nœud et re-triangulation non unique.

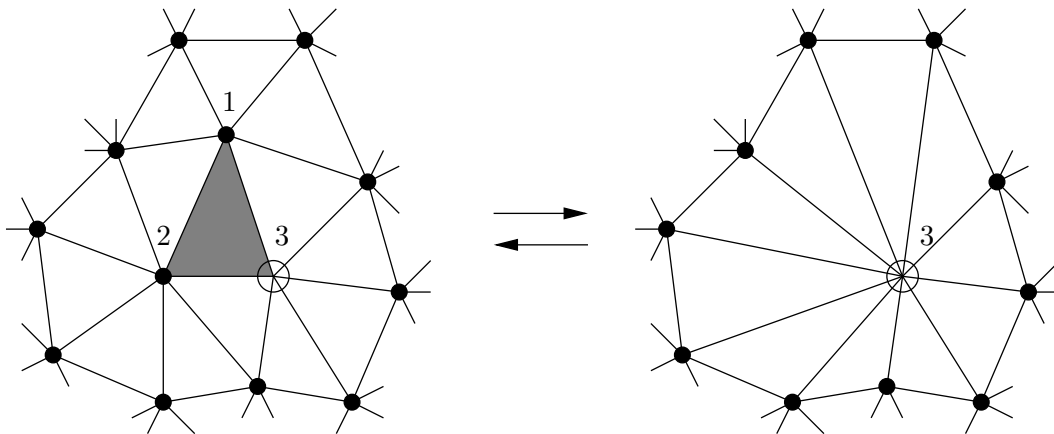
- **fusion parcimonieuse de deux nœuds**  $(1, 2) \rightarrow 1|2$  : elle ne conserve que l'un des deux sommets définissant une arête (voir figure 4.12).

FIG. 4.12 – Fusion parcimonieuse de deux nœuds :  $(1, 2) \rightarrow 2$  en haut ;  $(1, 2) \rightarrow 1$  en bas.

- **fusion optimisée de deux nœuds**  $(1, 2) \rightarrow 1'$  : elle agglomère les deux nœuds situés aux extrémités de l'arête  $(1, 2)$  (voir figure 4.13). La position du nœud résultant peut être optimisée.

FIG. 4.13 – Fusion optimisée de deux nœuds :  $(1, 2) \rightarrow 1'$ .

- **fusion parcimonieuse de trois nœuds**  $(1, 2, 3) \rightarrow 1|2|3$  : elle ne conserve que l'un des trois sommets définissant la face (voir figure 4.14).

FIG. 4.14 – Fusion parcimonieuse de trois nœuds :  $(1, 2, 3) \rightarrow 3$ .

- **fusion optimisée de trois nœuds**  $(1, 2, 3) \rightarrow 1'$  : elle agglomère les trois nœuds situés aux sommets de la face  $(1, 2, 3)$ . La position du nœud résultant peut être optimisée.
- **inversion d'arête** dans un quadrilatère convexe : elle est utilisée pour rendre un maillage plus uniforme, homogène, compact (voir figure 4.15). Dans le cadre du traitement numérique des images, elle permet aussi d'éviter l'effet de crénelage le long des contours des objets. Cette inversion n'est évidemment utilisée que si le quadrilatère est convexe. En effet, s'il est concave, l'inversion d'arête génère un maillage non conforme.

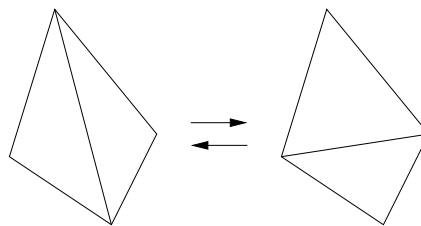


FIG. 4.15 – Inversion d'arête dans un quadrilatère convexe.

### Remarques

- 1) Ces opérateurs de décimation laissent le maillage conforme.
- 2) Ces opérateurs sont inversibles.

### B) La subdivision

Par construction (ajout de composants (nœud, arête, face)), la subdivision complexifie le maillage. Si elle n'est pas appliquée à toutes les mailles, le maillage obtenu peut être non conforme.

Si on note «  $1 \rightarrow n$  » le fait qu'un triangle donne naissance à  $n$  triangles, les subdivisions les plus couramment utilisées sont les subdivisions «  $1 \rightarrow 2$  », «  $1 \rightarrow 3$  » et «  $1 \rightarrow 4$  ». La figure 4.16 illustre de telles subdivisions.

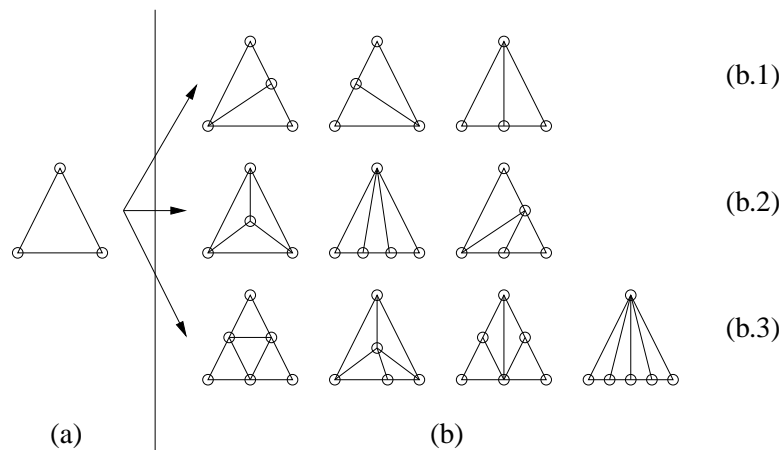


FIG. 4.16 – Exemples de subdivisions triangulaires. (a) : triangle initial ; (b) : subdivisions (b.1) : «  $1 \rightarrow 2$  », (b.2) : «  $1 \rightarrow 3$  » et (b.3) : «  $1 \rightarrow 4$  ».

On peut construire une hiérarchie de maillages en appliquant des subdivisions successives à partir du niveau le plus grossier.

### 4.1.3 Interpolation

Par construction, seuls les nœuds d'un maillage sont affectés d'une valeur (nodale) par l'intermédiaire d'une fonction  $v : \mathcal{S} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ . On étend  $v$  à  $\Omega \subset \mathbb{R}^2$  par interpolation. Le développement qui suit décrit en détail cette opération.

#### 4.1.3.1 Définitions et propriétés

##### 4.1.3.1.1 Définition (Polynôme d'interpolation de type Lagrange) (d'une fonction réelle d'une variable réelle)

Soient :

- $f : \mathbb{R} \rightarrow \mathbb{R}$ ,
- $\{x_i\}_{i=0..n}$ ,  $(n + 1)$  abscisses distinctes.

Un polynôme réel  $P_n^L$  à une indéterminée ( $\in \mathbb{R}[x]$ ) satisfaisant à :

$$P_n^L(x_i) = f(x_i), \forall i \in [0, n]_{\mathbb{N}},$$

est appelé **polynôme d'interpolation de type Lagrange** de la fonction  $f$ . Autrement dit, la fonction polynomiale  $x \mapsto P_n^L(x)$  passe exactement par tous les points  $(x_i, f(x_i))$ .

D'après le théorème de Lagrange, il existe un **unique** polynôme  $P_n^L$  de degré inférieur ou égal à  $n$  satisfaisant à la définition précédente. C'est le polynôme d'interpolation de Lagrange, à savoir :

$$\forall x \in \mathbb{R}, P_n^L(x) = \sum_{i=0}^n L_i(x) f(x_i) \quad (4.13)$$

avec :

$$L_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}. \quad (4.14)$$

##### 4.1.3.1.2 Définition (Polynôme d'interpolation de type Lagrange) (d'une fonction réelle de deux variables réelles)

Soient :

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,
- $\{(x_i, y_i)\}_{i=0..(m+1)(n+1)-1}$ ,  $(m + 1)(n + 1)$  points distincts de l'espace affine  $(\mathcal{E}_2, (O, \vec{i}, \vec{j}))$  rapporté à un repère orthonormé  $(O, \vec{i}, \vec{j})$ .

Un polynôme réel  $P_{(m+n)}^L$  à deux indéterminées ( $\in \mathbb{R}[x, y]$ ) de degré total au plus égal à  $m + n$  et de degrés partiels relativement à  $x$  au plus  $n$  et relativement à  $y$  au plus  $m$  :

$$P_{(m+n)}^L(x_i, y_i) = f(x_i, y_i), \forall i \in [0, (m + 1)(n + 1) - 1]_{\mathbb{N}},$$

est appelé **polynôme d'interpolation de type Lagrange** de la fonction  $f$ . Autrement dit, la fonction polynomiale  $(x, y) \mapsto P_{(m+n)}^L(x, y)$  passe exactement par tous les points  $(x_i, y_i, f(x_i, y_i))$ .

Le théorème de Lagrange se généralise en deux dimensions ainsi :

Soient  $\begin{cases} X = \{x_j \in \mathbb{R}\}_{j=0..n} \\ Y = \{y_i \in \mathbb{R}\}_{i=0..m} \end{cases}$ , alors il existe un **unique** polynôme  $P_{(m,n)}^L$  de degré  $m + n$  satisfaisant à la définition précédente pour tous les points de la grille rectangulaire  $XY \stackrel{n}{=} X \times Y$ , à savoir :

$$\forall (x, y) \in \mathbb{R}^2, P_{(m,n)}^L(x, y) = \sum_{i=0}^m \sum_{j=0}^n L_{i,j}(x, y) f(x_j, y_i) \quad (4.15)$$



avec :

$$L_{i,j}(x, y) = \prod_{k=0, k \neq j}^n \frac{x - x_k}{x_j - x_k} \prod_{l=0, l \neq i}^m \frac{y - y_l}{y_i - y_l} \quad (4.16)$$

---

### Remarque

Le polynôme d'interpolation couramment dénommée « bilinéaire » n'est rien d'autre que  $P_{(1,1)}^L$ .

---

#### 4.1.3.1.3 Définition (Polynôme d'interpolation de type Hermite) (d'une fonction réelle d'une variable réelle)

Soient :

- $f : \mathbb{R} \rightarrow \mathbb{R}$ ,
- $\{x_i\}_{i \in I_0}$ ,  $|I_0|$  abscisses distinctes.

Un polynôme réel  $P_n^H$  à une indéterminée ( $\in \mathbb{R}[x]$ ) satisfaisant à :

$$\begin{cases} P_n^H(x_i) = f(x_i), & \forall i \in I_0, \\ (P_n^H)'(x_i) = f^{(1)}(x_i), & \forall i \in I_1 \subset I_0, \\ \vdots \\ (P_n^H)^{(n)}(x_i) = f^{(n)}(x_i), & \forall i \in I_n \subset I_0, \end{cases}$$

où :

$$\sum_{i=0}^n |I_i| = n + 1,$$

est appelé **polynôme d'interpolation de type Hermite** de la fonction  $f$ . Autrement dit, la fonction polynomiale  $x \mapsto P_n^H(x)$  passe exactement par tous les points  $(x_i, f(x_i))$  et certaines dérivées de  $P_n^H$  sont contraintes à être égales à celles de  $f$ .

Notons que la généralisation de l'interpolation de Hermite en dimension deux se fait logiquement.

---

### Remarques

- 1) L'interpolation de type Lagrange, ne faisant pas intervenir de dérivée, s'oppose à l'interpolation de type Hermite qui fait intervenir des dérivées.
  - 2) Le polynôme interpolateur de type Hermite est plus régulier (plus lisse) que celui de Lagrange.
- 

#### 4.1.3.2 Application de l'interpolation au maillage

##### A) Cas des maillages triangulaires

On ne peut pas utiliser les polynômes généralisés  $P_{(m,n)}^L$  de Lagrange, car on ne dispose pas d'une grille rectangulaire de points, donc  $P_T^L$ , polynôme à deux indéterminées de degré total égal à 1, s'écrit :

$$\forall (x, y) \in \mathbb{R}^2, P_T^L(x, y) = a + bx + cy \quad (4.17)$$

et satisfait à :

$$\forall i \in \{1, 2, 3\}, P_T^L(x_i, y_i) = v(x_i, y_i) \stackrel{n}{=} v_i,$$

où  $(x_i, y_i), i \in \{1, 2, 3\}$  sont les trois sommets du triangle contenant  $(x, y)$ .

En appelant  $\beta = [a \ b \ c]^T$ , on obtient le système linéaire de trois équations à trois inconnues (a,b,c) :

$$A\beta = B,$$

où :

$$A = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}, \quad \beta = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} v(x_1, y_1) \\ v(x_2, y_2) \\ v(x_3, y_3) \end{bmatrix} \quad (4.18)$$

Les trois sommets n'étant pas alignés, le problème possède l'unique solution :

$$\beta = A^{-1}B.$$

La valeur interpolée en un point  $\mathbf{m} = (x, y)$  du triangle  $t$  est :

$$\forall \mathbf{m} = (x, y) \in t, P_T^L(x, y) = \sum_{i=1}^3 L_{T_i}(x, y)v(x_i, y_i) \quad (4.19)$$

où, cette fois-ci, les  $L_{T_i}$  se déduisent de la résolution précédente et sont donnés par :

$$\begin{bmatrix} L_{T_1}(x, y) \\ L_{T_2}(x, y) \\ L_{T_3}(x, y) \end{bmatrix} = A^{-T} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \quad (4.20)$$

$$\begin{bmatrix} L_{T_1}(x, y) \\ L_{T_2}(x, y) \\ L_{T_3}(x, y) \end{bmatrix} = \frac{1}{\text{Det}(A)} \begin{bmatrix} x_2y_3 - x_3y_2 & y_2 - y_3 & x_3 - x_2 \\ x_3y_1 - x_1y_3 & y_3 - y_1 & x_1 - x_3 \\ x_1y_2 - x_2y_1 & y_1 - y_2 & x_2 - x_1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \quad (4.21)$$

## B) Cas des maillages rectangulaires conformes

On peut utiliser les polynômes généralisés de Lagrange, car on dispose d'une grille de points, donc :

$$P_Q^L = P_{(1,1)}^L. \quad (4.22)$$

## C) À propos de l'interpolation de Hermite

Dans sa thèse [Dem02], Laurent Demaret a comparé les deux types d'interpolation sur des maillages triangulaires.

Il montre que l'interpolation de type Lagrange possède de mauvaises propriétés de différentiation sur les arêtes du maillage. En effet, une condition de différentiabilité à l'ordre 1 le long de l'arête commune à deux triangles est que les quatre points  $(x, y, I(x, y))$  soient coplanaires, ce qui se produit très rarement. Ainsi, lors de la reconstruction de l'image, la non-différentiabilité se traduit par l'apparition de blocs triangulaires.

Le choix de l'interpolation de type Hermite permet de réduire cet artefact. En effet, avec une interpolation de Hermite de degré total égal à deux et de degrés partiels égaux à deux obtenue en contraignant le gradient aux nœuds de l'arête (et les valeurs de la fonction en ces deux nœuds seuls, car on a ainsi 6 équations et 6 inconnues), on obtient une

différentiabilité d'ordre 1 le long de l'arête, ce qui est logique puisque le troisième point n'intervient pas. Pour rencontrer le problème précédent, il faut donc aller plus loin avec plus de conditions et un polynôme d'ordre supérieur.

Cependant, Hervé Le Floch [Flo97] a montré qu'en augmentant le degré des polynômes d'interpolation, la qualité du rendu finit par se dégrader et un coût de calcul important est ajouté. Ce phénomène de dégradation peut s'expliquer par le phénomène de Runge, présent en interpolation, qui a montré que le polynôme peut osciller avec des rebonds très prononcés.

#### D) Ce que nous en retenons

Dans cette thèse, nous traitons uniquement des maillages triangulaires et utilisons donc  $P_T^I$  (équation (4.17)) et la formulation de l'équation (4.21) pour l'implémentation.

#### 4.1.4 Les coordonnées barycentriques

Notre travail fait la part belle aux :

- coordonnées barycentriques (par exemple pour calculer les  $L_{T_i}$  de l'équation (4.21)),
- coordonnées barycentriques généralisées (en particulier dans notre algorithme de suivi d'objets déformables).

En voici les fondements essentiels.

##### 4.1.4.1 Choix de l'espace affine

On considère le plan affine  $\mathcal{E}_2$ , rapporté à un repère affine (cartésien)  $(O, (\vec{i}, \vec{j}))$ .

##### 4.1.4.2 Choix de la base affine dans $(\mathcal{E}_2, (O, (\vec{i}, \vec{j})))$

On sait (voir Ramis [[RDO01b], §5.2.3.2, pages 137–138]) que tout triplet de points  $(a_1, a_2, a_3) \in \mathcal{E}_2^3$  **non alignés** constitue une base affine (ou repère barycentrique) de l'espace affine  $\mathcal{E}_2$ .

##### 4.1.4.3 Définition des coordonnées barycentriques canoniques dans $(\mathcal{E}_2, (O, (\vec{i}, \vec{j})))$

###### 4.1.4.3.1 Définition (Coordonnées barycentriques dans $(\mathcal{E}_2, (O, (\vec{i}, \vec{j})))$ )

Étant donnés :

- $(a_1, a_2, a_3) \in \mathcal{E}_2^3$  *trois points non alignés*
- $\mathbf{m} \in \mathcal{E}_2$ ,

on appelle famille de **coordonnées barycentriques** de  $\mathbf{m}$  dans la base  $(a_i)_{i \in \{1,2,3\}}$  toute famille  $(\omega_i)_{i \in \{1,2,3\}} \in \mathbb{R}^3$  ( $\sum_{i=1}^3 \omega_i \neq 0$ ) vérifiant :

$$\overrightarrow{O\mathbf{m}} = \frac{1}{\sum_{i=1}^3 \omega_i} \sum_{i=1}^3 \omega_i \overrightarrow{Oa_i}. \quad (4.23)$$

---

**Remarques**

- 1) On démontre (voir Ramis [[RDO01b], §5.2.3 4, page 138]) que les  $(\omega_i)_{i \in \{1,2,3\}}$  sont définis à une constante multiplicative près.
  - 2)  $\mathbf{m}$  est le barycentre des points massiques  $(a_i, \omega_i)_{i \in \{1,2,3\}}$ , ce qui explique l'appellation.
  - 3) En imposant  $\sum_{i=1}^3 \omega_i = 1$ , alors la famille est unique et on parle de **coordonnées barycentriques canoniques**. Nous nous plaçons **toujours** dans cette situation.
- 

**4.1.4.4 Correspondance entre coordonnées cartésiennes de  $(\mathcal{E}_2, (O, (\vec{e}_1, \vec{e}_2)))$  et coordonnées barycentriques canoniques de  $(\mathcal{E}_2, (a_i)_{i=1,2,3})$** 

Supposons :

$$\begin{cases} \overrightarrow{O\mathbf{m}} = \sum_{i=1}^2 x_i \vec{e}_i = \sum_{j=1}^3 \omega_j \overrightarrow{Oa_j} \\ \overrightarrow{Oa_j} = \sum_{i=1}^2 a_{ji} \vec{e}_i \\ \sum_{j=1}^3 \omega_j = 1 \end{cases} \quad (4.24)$$

alors :

$$\begin{cases} x_i = \sum_{j=1}^3 \omega_j a_{ji} \\ \omega_j = L_{T_j}(x_1, x_2) \text{ (cf. équation (4.21))} \end{cases} \quad (4.25)$$

Notre implémentation utilise l'équation (4.21) pour calculer aisément aussi bien les coordonnées barycentriques d'un point  $\mathbf{m}$  que les coefficients du polynôme interpolateur  $P_T^L$ . En effet, la complexité algorithmique est faible :

- précalcul de la matrice  $3 \times 3$  ( $A^{-T}$ ) de l'équation (4.21) pour chaque triangle,
- produit « matrice-vecteur ».

**4.1.4.5 Définition des coordonnées barycentriques généralisées dans  $(\mathcal{E}_2, (O, (\vec{i}, \vec{j})))$** 
**A) Définition**
**4.1.4.5.1 Définition (Coordonnées barycentriques généralisées dans  $(\mathcal{E}_2, (O, (\vec{i}, \vec{j})))$ )**

Soient  $a_1, \dots, a_k$ ,  $k(> 3)$  points (trois d'entre eux au moins  $(a_{i_1}, a_{i_2}, a_{i_3})$  n'étant pas alignés) de  $\mathcal{E}_2$  alors,  $\forall \mathbf{m} \in \mathcal{E}_2$ ,  $\exists (\omega_1, \dots, \omega_k) \in \mathbb{R}^k$  non uniques tels que :

$$\overrightarrow{O\mathbf{m}} = \frac{1}{\sum_{i=1}^k \omega_i} \sum_{i=1}^k \omega_i \overrightarrow{Oa_i}. \quad (4.26)$$

Par abus de langage, les  $(\omega_i)_{i=1..k}$  sont appelées **coordonnées barycentriques généralisées (cbg)** du point  $\mathbf{m}$ .

---

**Remarques**

- 1) L'existence est assurée à partir de la définition 4.1.4.3.1 appliquée à la base affine  $(a_{i_1}, a_{i_2}, a_{i_3})$ .

- 2) Pour la négation de l'unicité, il suffit d'imaginer une autre base  $(a_{j_1}, a_{j_2}, a_{j_3})$  distincte de  $(a_{i_1}, a_{i_2}, a_{i_3})$  qui conduit à un deuxième triplet de coordonnées qui représente le même point.

## B) Quelques exemples

Les coordonnées barycentriques généralisées n'étant pas uniques, en voici une liste non exhaustive à partir des notations de la figure 4.17, où nous supposons que les  $a_i$  sont les sommets d'un polygone  $\Pi$ .

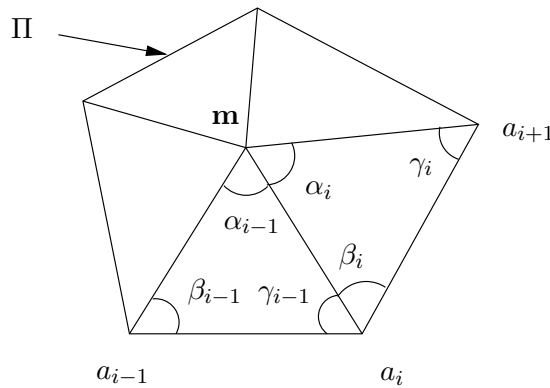


FIG. 4.17 – Notations pour les coordonnées barycentriques généralisées.

### B.1) Coordonnées de Wachspress [Fix78]

$$\omega_i = \frac{A(a_{i-1}, a_i, a_{i+1})}{A(a_{i-1}, a_i, \mathbf{m})A(a_i, a_{i+1}, \mathbf{m})}, \quad i \in [1, k]_{\mathbb{N}}, \quad (4.27)$$

où  $A(a_{i-1}, a_i, a_{i+1})$  représente l'aire du triangle  $\widehat{a_{i-1}a_i a_{i+1}}$  dont les sommets sont ordonnés  $(a_{i-1}, a_i, a_{i+1})$ .

### B.2) Coordonnées de Meyer *et al.* [Mey02]

Leur expression est équivalente à celle de Wachspress, à savoir [[Mey02], éq. (7), p. 15] :

$$\omega_i = \frac{\cotg(\gamma_{i-1}) + \cotg(\beta_i)}{\|\mathbf{m} - a_i\|^2}, \quad i \in [1, k]_{\mathbb{N}}. \quad (4.28)$$

### B.3) Coordonnées de Floater [Flo03]

Alors que tout barycentre est invariant par transformation affine, les systèmes de coordonnées proposés ne le sont pas tous : par exemple, les cbg de Floater, appelées « valeur moyenne », [FHK06], [[Flo03], page 21] ne le sont pas :

$$\omega_i = \frac{tg(\frac{\alpha_{i-1}}{2}) + tg(\frac{\alpha_i}{2})}{\|\mathbf{m} - a_i\|}, \quad i \in [1, k]_{\mathbb{N}}, \quad 0 < \alpha_i < \pi. \quad (4.29)$$

---

**Remarque**

Il est à noter que les cbg peuvent être négatives dans le cas où le polygone n'est pas convexe ( $\gamma_{i-1} + \beta_i > \pi$ ).

---

D'après le théorème 5.3.1 (3°) p. 140 de Ramis *et al.* [RDO01b], une fonction  $f : \mathcal{E}_2 \rightarrow \mathcal{E}_2$  est affine si et seulement si elle conserve le barycentre, c'est-à-dire, si et seulement si

$$\forall (a_i, \omega_i)_{i \in I}, \overrightarrow{Og} = \sum_{i \in I} \omega_i \overrightarrow{Oa_i} \Leftrightarrow f(\overrightarrow{Og}) = \sum_{i \in I} \omega_i f(\overrightarrow{Oa_i}),$$

**C) Ce que nous en retenons**

Dans ce travail de thèse, nous choisissons les coordonnées de Wachspress [Fix78, Mey02] pour leur propriété d'invariance par transformation affine. Par contre, elles ne sont définies que pour des polygones convexes. Dans notre cas, lors de l'évolution du maillage, cette contrainte n'est pas toujours assurée. Actuellement, nous l'ignorons et utilisons tout de même les coordonnées barycentriques généralisées. Le seul risque auquel nous courons est d'aboutir à un échec du suivi ou d'augmenter le nombre d'itérations pour suivre...

**4.1.4.6 Utilisation des coordonnées barycentriques généralisées**

Les coordonnées barycentriques généralisées sont majoritairement utilisées pour :

- paramétrer les faces d'un objet ;
- interpoler sur un polygone d'ordre quelconque ;
- modéliser des surfaces ;
- approcher les surfaces de Bézier ;
- déformer « sans limites » [Mey02].

Nous avons eu l'initiative d'utiliser les coordonnées barycentriques généralisées en vision, plus précisément pour approcher la composition de deux transformations d'un maillage (voir §4.1.5.5).

**4.1.5 Transformations géométriques d'un maillage, utiles au suivi d'objets**

Dans cette thèse, l'outil mathématique maillage est principalement utilisé pour approcher, dans une séquence d'images, les déformations entre deux images consécutives. À une déformation d'une image est donc associée une déformation du maillage. Voici les déformations géométriques d'un maillage que nous prenons en compte :

- les transformations globales ;
- les transformations nodales ;
- les transformations nodales inverses ;
- la composition d'une transformation globale avec une transformation nodale.

**4.1.5.1 Choix des paramètres**

Nous ne traitons que les maillages triangulaires, alors :

- soit les paramètres  $\beta$  de ce modèle sont les  $6K$  composantes des transformations affines subies par les  $K$  triangles du maillage, à savoir :

$$\beta \in (\mathbb{R}^6)^K = \left[ \dots \beta^{t(j)\top} \dots \right]^\top, \quad (4.30)$$

où  $\beta^{t(j)} = [\beta_1^{t(j)} \dots \beta_6^{t(j)}]^\top$  définit la transformation affine associée au triangle  $j$ . Nakaya et Harashima [[NH94], §III.A, p. 342], d'une part, et Altunbasak et Tekalp [[AT97b], §II.B, p. 1256], d'autre part, utilisent ce jeu de paramètres.

- soit les paramètres  $\beta$  de ce modèle sont les  $2N$  composantes des translations subies par les  $N$  nœuds du maillage, à savoir :

$$\beta \in (\mathbb{R}^2)^N = \left[ \dots \beta^{n(i)\top} \dots \right]^\top, \quad (4.31)$$

où  $\beta^{n(i)} = [\beta_x^{n(i)} \beta_y^{n(i)}]^\top$  définit la translation du nœud  $S_i$ . Marquant *et al.* [MPL00] utilisent ce jeu de paramètres.

#### 4.1.5.2 Transformations globales

Une **transformation globale** du maillage triangulaire est définie à l'aide d'une transformation affine par morceaux, à savoir autant de transformations affines que de triangles. La figure 4.18 en est une illustration. Le maillage en traits pleins est transformé en le maillage en traits discontinus.

##### Remarque

Par construction, on peut définir une transformation globale inverse, elle sera elle aussi affine par morceaux, à savoir autant de transformations affines que de triangles en traits discontinus.

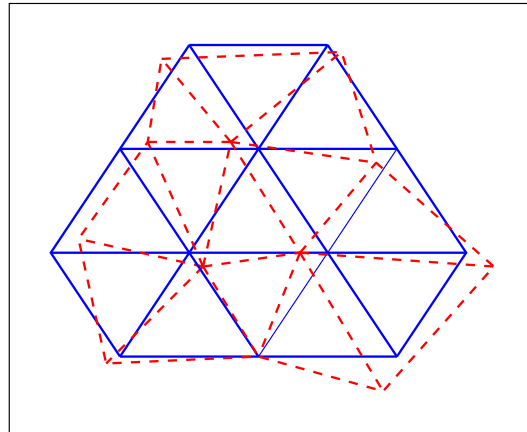


FIG. 4.18 – Exemple de transformation globale  $T$ . « — » maillage initial ; « - - » maillage déformé.

Les transformations globales doivent intégrer les contraintes de continuité entre deux triangles adjacents. Nous choisissons le second jeu de paramètres. Ce choix est d'autant plus justifié que toute perturbation des paramètres a une représentation physique immédiate et les contraintes de continuité sont intrinsèquement vérifiées. Le maillage déformé conserve sa

topologie. Notons que pour une transformation globale quelconque, le maillage peut devenir non conforme : nous ne sommes pas à l'abri d'un retournement de maille.

Seuls les maillages triangulaires sont traités ; aussi notons-nous en forme de triangle  $\Delta_i$  la face  $F_i$  du maillage initial et  $\Delta'_i$  la face correspondante dans le maillage déformé, à savoir :

$$\Delta'_i = T_{Aff_i}(\Delta_i, \beta), \quad (4.32)$$

où  $T_{Aff_i}$  est une transformation affine. Ses six paramètres effectifs se déduisent très facilement de  $\beta \in (\mathbb{R}^2)^N$  à partir des trois translations  $\beta^{n(j)} \in \mathbb{R}^2$  des trois sommets de  $\Delta_i$ . En rappelant que le domaine  $\Omega$  du maillage est tel que  $\Omega = \bigcup_{j \in J_K} \Delta_j$ , où  $J_K = [1, K]_{\mathbb{N}}$ , l'expression analytique de la transformation globale  $T$  est donc :

$$T : \bigcup_{j \in J_K} \Delta_j \subset \mathbb{R}^2 \times (\mathbb{R}^2)^N \rightarrow \bigcup_{j \in J_K} \Delta'_j$$

où :

$$\forall j \in J_K, \begin{cases} T|_{\Delta_j} : \Delta_j \times (\mathbb{R}^2)^N \rightarrow \Delta'_j \\ (\mathbf{m}, \beta) \mapsto \mathbf{m}' = T|_{\Delta_j}(\mathbf{m}, \beta) = T_{Aff_j}(\mathbf{m}, \beta). \end{cases} \quad (4.33)$$

#### 4.1.5.3 Transformations nodales / nodales inverses

Toute transformation globale du maillage peut se décomposer en une succession de transformations que nous avons choisi d'appeler **transformations nodales**  $T_{\beta^{(n)}}$  qui sont caractérisées par le déplacement  $\beta^{(n)} \in \mathbb{R}^2$  d'un seul nœud  $S_n$ . La frontière du polygone support  $\Pi_n$  de  $S_n$  est donc invariante par transformation nodale. Il en est de même pour toutes les mailles n'appartenant pas à  $\Pi_n$ . Sur la figure 4.19, le nœud  $S_n$  est au centre du maillage.

Ces transformations sont :

- les plus simples que l'on puisse appliquer à un maillage,
- inversibles.

Pour ces deux raisons, notre intérêt se porte sur elles et ceci justifie notre choix des paramètres de la transformation : un vecteur de translation par nœud.

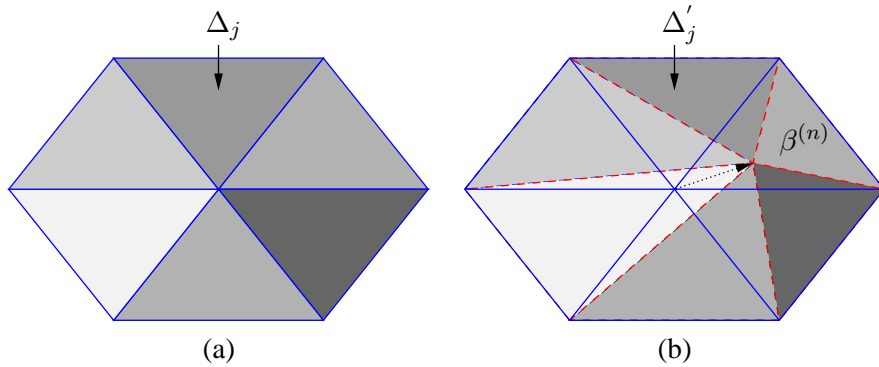


FIG. 4.19 – Exemple de transformation nodale  $T_{\beta^{(n)}}$ . (a) : en traits pleins, maillage initial  $\mathcal{M}$ ; (b) : en tirets, maillage initial ayant subi une transformation nodale, soit  $T_{\beta^{(n)}}(\mathcal{M})$ .

L'expression analytique de la transformation nodale  $T_{\beta^{(n)}}$  est donnée par (voir figure 4.19) :



$$\left\{ \begin{array}{lcl} T_{\beta^{(n)}} : & \bigcup_{j \in J_K} \Delta_j & \rightarrow \bigcup_{j \in J_K} \Delta'_j \\ & \mathbf{m} & \mapsto T_{\beta^{(n)}}(\mathbf{m}) = T(\mathbf{m}, \beta) \end{array} \right. \quad (4.34)$$

où  $\beta = [0_{1,2n-2} \quad \beta^{(n)T} \quad 0_{1,2N-2n}]^T$ . Autrement dit, en notant  $\Pi_n = \bigcup_{j \in J_n} \Delta_j$ , où  $J_n = \{j \in J_K \mid S_n \in \Delta_j\}$ , on a :

$$\left\{ \begin{array}{l} \forall j \in J_n, \quad \left\{ \begin{array}{lcl} T_{\beta|_{\Delta_j}^{(n)}} : & \Delta_j & \rightarrow \Delta'_j \\ & \mathbf{m} & \mapsto \mathbf{m}' = T_{\beta|_{\Delta_j}^{(n)}}(\mathbf{m}) = T_{Aff_j}(\mathbf{m}, \beta) \end{array} \right. \\ \text{et} \\ \forall j \in J_K \setminus J_n, \quad \left\{ \begin{array}{lcl} T_{\beta|_{\Delta_j}^{(n)}} : & \Delta_j & \rightarrow \Delta'_j \\ & \mathbf{m} & \mapsto \mathbf{m}' = T_{\beta|_{\Delta_j}^{(n)}}(\mathbf{m}) = \mathbf{m}. \end{array} \right. \end{array} \right. \quad (4.35)$$

Une transformation nodale est bijective de transformation nodale inverse  $T_{\beta^{(n)}}^{-1}$ , à savoir :

$$T_{\beta^{(n)}}^{-1} : \bigcup_{j \in J_K} \Delta'_j \rightarrow \bigcup_{j \in J_K} \Delta_j$$

où :

$$\left\{ \begin{array}{l} \forall j \in J_n, \quad \left\{ \begin{array}{lcl} T_{\beta|_{\Delta'_j}^{(n)}}^{-1} : & \Delta'_j & \rightarrow \Delta_j \\ & \mathbf{m} & \mapsto \mathbf{m}' = T_{\beta|_{\Delta'_j}^{(n)}}^{-1}(\mathbf{m}) = T_{Aff_j}^{-1}(\mathbf{m}, \beta) \end{array} \right. \\ \text{et} \\ \forall j \in J_K \setminus J_n, \quad \left\{ \begin{array}{lcl} T_{\beta|_{\Delta'_j}^{(n)}}^{-1} : & \Delta'_j & \rightarrow \Delta_j \\ & \mathbf{m} & \mapsto \mathbf{m}' = T_{\beta|_{\Delta'_j}^{(n)}}^{-1}(\mathbf{m}) = \mathbf{m}. \end{array} \right. \end{array} \right. \quad (4.36)$$

---

#### Remarque

Avec le second jeu de paramètres  $\beta \in (\mathbb{R}^2)^N$ , une transformation nodale est paramétrée avec uniquement deux paramètres  $\beta^{(n)} \in \mathbb{R}^2$  non nuls.

Avec le premier jeu de paramètres  $\beta \in (\mathbb{R}^6)^K$ , la même transformation nodale est paramétrée par  $6|J_n|$  paramètres non nuls.

Avec le premier jeu de paramètres  $\beta \in (\mathbb{R}^6)^K$ , il y a redondance d'information due au fait qu'un sommet appartient à plusieurs triangles. Ainsi des paramètres peuvent se déduire des autres.

---

#### 4.1.5.4 Application d'une transformation nodale inverse à un maillage dépendant de l'original

Dans notre algorithme de suivi (voir §4.3.1.3), la transformation nodale inverse est incontournable. Caractérisons cette transformation, autrement dit, calculons géométriquement

l'image d'un maillage  $\mathcal{M}^q$  de domaine  $\Omega^q \subset \Omega$  (domaine de  $\mathcal{M}$ ) par une transformation nodale inverse.

On obtient  $\mathcal{M}^{qT_{\beta(n)}^{-1}} = T_{\beta(n)}^{-1}(\mathcal{M}^q)$  en cherchant l'image de chaque arête de  $\mathcal{M}^q$  par  $T_{\beta(n)}^{-1}$ . Étant donné qu'un segment a pour image un segment par transformation affine, on cherche l'image de ses deux extrémités.

La figure 4.20 illustre l'application d'une telle transformation sur un maillage « quelconque »  $\mathcal{M}^q = T_{\beta(q(n))}(\mathcal{M})$  :

- en bleu (traits pleins), le maillage initial  $\mathcal{M}$ ,
- en rouge (tirets),  $\mathcal{M}^{T_{\beta(n)}} = T_{\beta(n)}(\mathcal{M})$ ,
- en vert (point-tiret),  $\mathcal{M}^q$ ,
- en magenta (en pointillé),  $\mathcal{M}^{qT_{\beta(n)}^{-1}} = T_{\beta(n)}^{-1}(\mathcal{M}^q)$ , le maillage recherché.

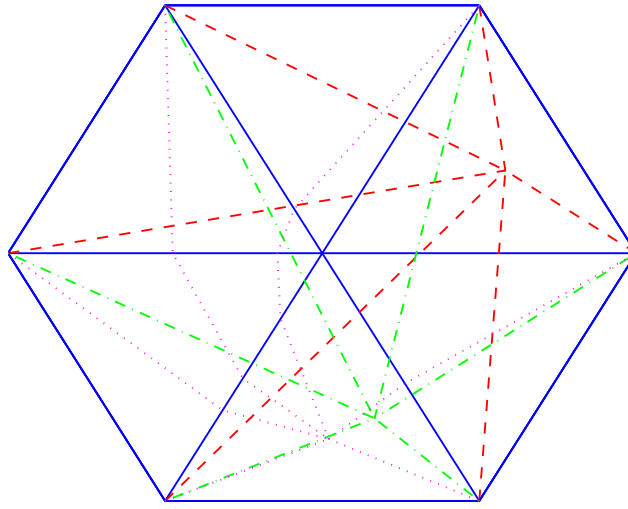


FIG. 4.20 – Transformation nodale inverse appliquée à un maillage « quelconque »  $\mathcal{M}^q$  : en bleu (traits pleins), le maillage initial  $\mathcal{M}$  ; en rouge (tirets),  $\mathcal{M}^{T_{\beta(n)}}$  ; en vert (point-tiret),  $\mathcal{M}^q$  ; en magenta (pointillé),  $\mathcal{M}^{qT_{\beta(n)}^{-1}}$

L'algorithme 4.1.3 fournit un cadre générique à la solution géométrique qui est détaillée dans l'algorithme 4.1.4.

## A) Algorithme générique

**Pour** chaque arête ( $a$ ) de  $\mathcal{M}^q$  **Faire**

**Pour**  $i$  de 1 à  $K_n$  **Faire**

$a_i (\stackrel{n}{=} [e'_1 e'_2]) \leftarrow a \cap \Delta'_i$

**Si**  $a_i \neq \emptyset$  **Alors**

$e_1 \leftarrow g_{th}(e'_1, i, \mathcal{M}, \mathcal{M}^{T_{\beta(n)}})$

$e_2 \leftarrow g_{th}(e'_2, i, \mathcal{M}, \mathcal{M}^{T_{\beta(n)}})$

      tracer le segment  $[e_1, e_2]$

**Fin Si**

**Fin Pour**

**Fin Pour**

Fonction utile :

$e' = g_{th}(e, k, \mathcal{M}, \mathcal{M}')$  : construction géométrique (détails dans l'algorithme 4.1.4)

Algorithme 4.1.3: Algorithme de construction géométrique de la transformée nodale inverse d'un maillage  $\mathcal{M}^q$ .

## B) Algorithme de construction géométrique

## B.1) Les composants géométriques du problème

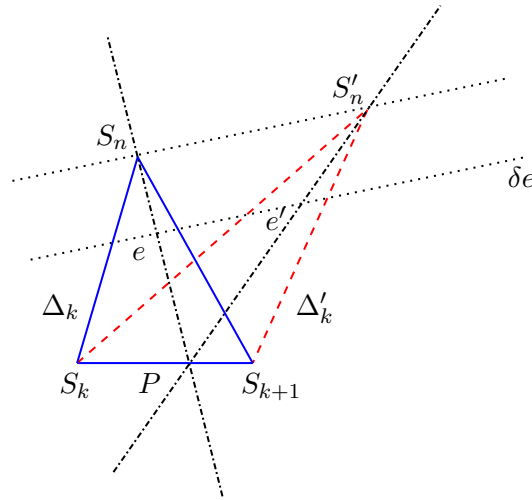


FIG. 4.21 – Construction géométrique de l'image du point  $e' \in \Delta'_k$  par la transformation affine inverse qui transforme le triangle  $\Delta'_k$  (rouge en tirets) en le triangle  $\Delta_k$  (bleu en traits pleins).

## B.2) Les équations

Soient :

-  $T_{Aff_k}$  l'application affine telle que :

- \*  $S_n \xrightarrow{T_{Aff_k}} S'_n \xrightarrow{T_{Aff_k}^{-1}} S_n,$
- \*  $S_k \xrightarrow{T_{Aff_k}} S_k \xrightarrow{T_{Aff_k}^{-1}} S_k,$
- \*  $S_{k+1} \xrightarrow{T_{Aff_k}} S_{k+1} \xrightarrow{T_{Aff_k}^{-1}} S_{k+1},$
- $P = (S'_n e') \cap (S_k S_{k+1})$

alors :

$$(S'_n P) \xrightarrow{T_{Aff_k}^{-1}} (S_n P)$$

$e' \in (S'_n P)$  est barycentre des points massiques  $(S'_n, \overline{e'P})$  et  $(P, \overline{e'S'_n})$ .

Or toute transformation affine conserve le barycentre, donc  $e = T_{Aff_k}^{-1}(e') \in (S_n P)$  est barycentre des points massiques  $(S_n, h' \overline{e'P})$  et  $(P, h' \overline{e'S'_n})$ .

De plus,  $e \in (S_n P)$  est barycentre des points massiques  $(S_n, \overline{eP})$  et  $(P, \overline{eS_n})$ .

On a donc  $h' \overline{e'P} = h \overline{eP}$  et  $h' \overline{e'S'_n} = h \overline{eS_n}$  d'où :

$$\frac{\overline{e'P}}{\overline{e'S'_n}} = \frac{\overline{eP}}{\overline{eS_n}}.$$

On note  $(\delta e)$  la droite  $(ee')$ . En appliquant la réciproque du théorème de Thalès, on obtient le parallélisme des droites  $(\delta e)$  et  $(S_n S'_n)$ .

D'où l'algorithme 4.1.4.

### B.3) L'algorithme

<b>fonction</b> $e' = g_{th}(e, k, \mathcal{M}, \mathcal{M}')$
$P \leftarrow (S_k S_{k+1}) \cap (S'_n e)$
Calculer la droite $(\delta e)$ parallèle à $(S_n S'_n)$ , passant par $e$
$e' \leftarrow (\delta e) \cap (S_n P)$
<b>Renvoyer</b> $e'$

Algorithme 4.1.4: Algorithme de construction géométrique de l'image d'un point par une transformation affine inverse utilisant le théorème de Thalès.

### C) Ce que nous en retenons

- i) Sur la figure 4.20, certaines arêtes sont transformées en des lignes brisées : ce n'est donc plus un maillage au sens de nos définitions.

L'explication découle logiquement de l'action « **Pour**  $i$  **de** 1 **à**  $K_n$  » de l'algorithme 4.1.3 qui signifie qu'une arête  $a$  est tronçonnée en autant de segments que de triangles  $\Delta'_i$  traversés.

- 2i) Dans notre algorithme de suivi,  $\mathcal{M}^q = \mathcal{M}$ , c'est ce que la figure 4.22 illustre.

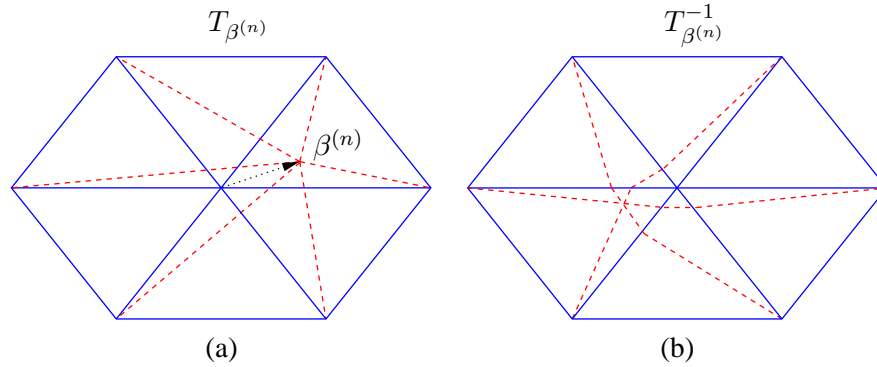


FIG. 4.22 – En traits pleins, le maillage initial  $\mathcal{M}$ . En traits discontinus : (a) la transformée par  $T_{\beta^{(n)}}$  du maillage initial ; (b) la transformée par  $T_{\beta^{(n)}}^{-1}$  du maillage initial.

- 3i) Notons que les caractérisations exactes, analytique et géométrique, d'une transformation nodale inverse n'ont jamais été utilisées en vision : c'est l'une de nos contributions dans ce domaine.

#### 4.1.5.5 Compositions exacte et approchée d'une transformation nodale et d'une transformation globale définies sur le même maillage

Notre algorithme de suivi d'objets déformables (représentés par des maillages) (voir §4.3.1.3) repose sur des transformations obtenues par composition d'une transformation nodale et d'une transformation globale définies sur le même maillage initial. Le développement qui suit s'articule autour de :

- la composition exacte et ses difficultés intrinsèques ;
- des leviers potentiels d'une composition approchée ;
- la composition approchée de Matthews et Baker ;
- notre composition approchée.

#### A) La composition exacte ( $T^{gon}$ ) et ses difficultés intrinsèques

##### A.1) Les équations

Soient :

- $\mathcal{M}$  le maillage initial (voir figure 4.23-(a)),
- $\mathcal{M}_{T_{\beta^{(n)}}} = T_{\beta^{(n)}}(\mathcal{M})$ , le maillage  $\mathcal{M}$  déformé par la transformation nodale  $T_{\beta^{(n)}}$  (voir figure 4.23-(b)),
- $\mathcal{M}_T = T(\mathcal{M}, \beta)$  (voir équation (4.33)), le maillage  $\mathcal{M}$  déformé par la transformation globale  $T$  (voir figure 4.23-(c)),

on cherche à caractériser la transformation  $T^{gon} = T \circ T_{\beta^{(n)}}$  définie par :

$$\begin{cases} T^{gon} : \mathbb{R}^2 \times \mathbb{R}^p & \rightarrow \mathbb{R}^2 \\ (\mathbf{m}, \beta) & \mapsto T^{gon}(\mathbf{m}, \beta) = T(T_{\beta^{(n)}}(\mathbf{m}), \beta). \end{cases} \quad (4.37)$$

La figure 4.23-(d) illustre l'application d'une telle transformation sur le maillage  $\mathcal{M}$ .

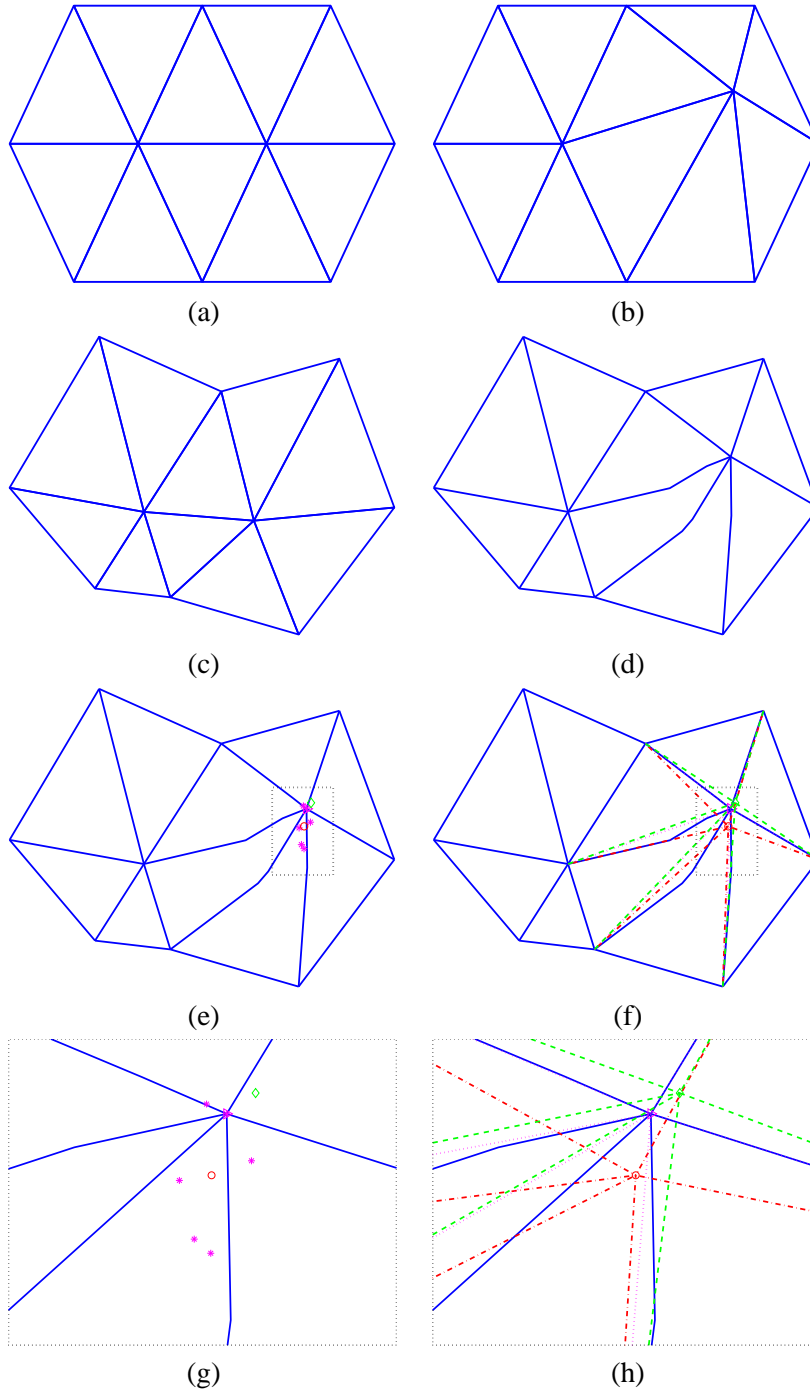


FIG. 4.23 – Illustration de la composition de transformations globale et nodale de maillage : (a) le maillage initial  $\mathcal{M}$ ; (b)  $\mathcal{M}_{T_{\beta(n)}} = T_{\beta(n)}(\mathcal{M})$ ; (c)  $\mathcal{M}_T = T(\mathcal{M}, \beta)$ ; (d)  $\mathcal{M}' = T(T_{\beta(n)}(\mathcal{M}), \beta) = T^{gon}(\mathcal{M})$ ; (e) et (f) propositions pour  $\tilde{T}^{gon}$ , (g) et (h) zooms sur la zone détournée par le rectangle en pointillé noir des figures (e) et (f). Les différentes approximations pour  $S''_n$  sont respectivement représentées par \* (magenta) pour B-(i),  $\triangleright$  (magenta) pour B-(2i),  $\circ$  (rouge) pour Matthews et Baker et  $\diamond$  (vert) pour nous. Les différentes approximations pour  $T^{gon}$  sont respectivement représentées par des pointillés ' . ' magenta pour B-(i) et B-(2i), des points-tirets ' . - ' rouges pour Matthews et Baker et des tirets ' - - ' verts pour nous.

### A.2) Les difficultés

Cette définition de  $T^{gon}$  soulève au moins deux difficultés :

- i) sur le plan théorique, que se passe-t-il lorsque l'inclusion :

$$T_{\beta(n)}(\mathcal{M}) = \mathcal{M}_{T_{\beta(n)}} \subset \bigcup_{i=1..K} \Delta_i$$

n'est pas satisfaite ? Comment étendre  $T$  en pratique ?

- 2i) l'exemple des lignes brisées apparaissant lors de l'application d'une transformation nodale inverse (voir figure 4.20) n'est qu'un cas particulier de résultats inexploitable en sortie d'une transformation globale  $T$  appliquée à un maillage  $\mathcal{M}' \neq \mathcal{M}$ . La figure 4.23-(d) constitue un second contre-exemple pour lequel  $T$  est simplement une transformation affine par morceaux.

On peut énoncer les propriétés suivantes :

- Une première condition suffisante de génération de lignes brisées dans  $\mathcal{M}' = T(\mathcal{M}_{T_{\beta(n)}}, \beta)$  (ce serait le cas pour tout  $T(\mathcal{M}_q, \beta)$ ) est :

- 1) l'existence d'une maille  $\Delta'_i \in \mathcal{M}'$  telle que :

$$\forall \Delta_j \in \mathcal{M}, \Delta'_i \cap \Delta_j \neq \Delta'_i$$

(autrement dit : aucune maille  $\Delta_j$  ne contient  $\Delta'_i$ ),

- 2)  $T$  n'est pas affine globale sur  $\bigcup_{j \in J} \Delta_j$ , où  $J = \{j \in J_K | \Delta'_i \cap \Delta_j \neq \emptyset\}$ .

- Une condition suffisante pour satisfaire la clause (1) précédente est  $T_{\beta(n)} \neq Id$ .

### A.3) Ce que nous en retenons

En dehors du cas où  $T$  est une transformation affine globale de  $\mathcal{M}$ ,  $T^{gon}$  n'est pas une transformation globale de  $\mathcal{M}$ .

### B) La composition approchée ( $\tilde{T}^{gon}$ )

#### B.1) Des leviers potentiels d'une composition approchée

Les difficultés intrinsèques à  $T^{gon}$  conduisent à remplacer  $T$  par  $\tilde{T}$  et, en conséquence,  $T^{gon}$  par une transformation globale de  $\mathcal{M}$ ,  $\tilde{T}^{gon}$ , qui conserve la topologie de  $\mathcal{M}$  sans briser les arêtes.

Mathématiquement,  $\tilde{T}^{gon}$  satisfait à :

$$\tilde{T}^{gon}(S_i, \tilde{\beta}) = T^{gon}(S_i, \beta), \forall S_i \in \mathcal{S} \setminus \{S_n\}, \quad (4.38)$$

où :

$$\forall j \in [1, 2N]_{\mathbb{N}} \setminus \{2n-1, 2n\}, \tilde{\beta}_j = \beta_j.$$

Il reste à définir  $\{\beta_{2n-1}, \beta_{2n}\}$  ou encore  $\tilde{T}^{gon}(S_n, \tilde{\beta})$ .

En notant  $\Pi_n$  le polygone support de  $S_n$ ,  $\Pi_n = \bigcup_{j \in J_n} \Delta_j$ , où  $J_n = \{j \in J_K | S_n \in \Delta_j\}$ , et  $S'_n = T_{\beta(n)}(S_n)$ , mentionnons quelques pistes qui tombent sous le sens :

- i)  $S''_n = T_{|\Delta_i|}(S'_n)$ , où  $i \in J_n$  est choisi « au hasard », d'où  $[\tilde{\beta}_{2n-1} \tilde{\beta}_{2n}]^T = S''_n - S_n$ .  
Il y a  $|J_n|$  possibilités (voir les figures 4.23-(e) et (g), où les  $S''_n$  sont les \* magenta)
- 2i)  $S''_n = T_{|\Delta_{i_0}|}(S'_n)$ , où  $\begin{cases} i_0 \in J_n \\ S'_n \in \Delta_{i_0} \end{cases}$ , d'où  $[\tilde{\beta}_{2n-1} \tilde{\beta}_{2n}]^T = S''_n - S_n$  (voir les figures 4.23-(e) et (g), où  $S''_n$  est le  $\triangleright$  magenta et les figures 4.23-(f) et (h), où  $\tilde{T}^{gon}$  est le maillage magenta en pointillé ' . ').

### Remarques

- 1) On ne traite toujours pas la difficulté A.2-(i).
- 2) Dans le cas où  $T$  est une transformation affine globale, toutes les solutions approchées sont exactes.

### B.2) La composition approchée de Matthews et Baker

Ils remplacent les deux propositions précédentes par [[MB04], §4.1.4 , p. 147] :

$$S''_n = \frac{1}{|J_n|} \sum_{j \in J_n} T_{|\Delta_j|}(S'_n), \quad (4.39)$$

d'où  $[\tilde{\beta}_{2n-1} \tilde{\beta}_{2n}]^T = S''_n - S_n$  (voir les figures 4.23-(e) et (g), où  $S''_n$  est le  $\circ$  rouge et les figures 4.23-(f) et (h), où  $\tilde{T}^{gon}$  est le maillage rouge en points-tirets ' - ').

### Remarque

Les deux remarques de la section 4.1.5.5.B.1 s'appliquent.

### B.3) Notre composition approchée

Elle est fondée sur la propriété d'invariance des coordonnées barycentriques généralisées de Wachspress (formulation de Meyer, équation (4.28)) par transformation affine globale. Voici les équations qui conduisent à l'algorithme 4.1.5.

Soient :

- $(\lambda_j)_{j=1..|B|}$  les coordonnées barycentriques généralisées de  $S'_n$  dans la base  $B$  formée par les sommets de  $\Pi_n$ ,
- $B'' = T(B, \beta)$ ,

alors  $S''_n$  est défini par les coordonnées barycentriques généralisées  $(\lambda_j)_{j=1..|B|}$  dans la base  $B''$ , d'où  $[\tilde{\beta}_{2n-1} \tilde{\beta}_{2n}]^T = S''_n - S_n$  (voir les figures 4.23-(e) et (g), où  $S''_n$  est le  $\diamond$  vert et les figures 4.23-(f) et (h), où  $\tilde{T}^{gon}$  est le maillage vert en tirets ' - ').



**fonction**  $S_n'' = \text{approximation\_avec\_cbg}(\mathcal{M}, n, \beta^{(n)}, \beta)$   
 $J \leftarrow \text{polygone\_support\_vers\_sommets}(\mathcal{M}, n)$   
 $\mathcal{B} \leftarrow \{S_j\}_{j \in J}$   
 $S_n' \leftarrow T_{\beta^{(n)}}(S_n)$   
 $\lambda \leftarrow \text{cbg\_wachspress}(S_n', \mathcal{B}) \ (\lambda \in \mathbb{R}^{|J|})$   
 $S_n'' \leftarrow \sum_{j \in J} \lambda_j \underbrace{(S_j + [\beta_{2j-1} \ \beta_{2j}]^T)}_{T(S_j, \beta)}$

**Fonctions utiles :**

$J = \text{polygone\_support\_vers\_sommets}(\mathcal{M}, n)$  : renvoie les indices des nœuds de  $\mathcal{M}$ , ordonnés selon le sens trigonométrique, qui sont sommets de  $\Pi_n$ .

$\lambda = \text{cbg\_wachspress}(S, \mathcal{B})$  : renvoie les coordonnées barycentriques généralisées de  $S$  dans la base  $\mathcal{B}$  (Wachspress/Meyer, équation (4.28)).

Algorithme 4.1.5: Algorithme d'approximation de la transformation  $T^{gon}$  : notre proposition.

### Remarques

- 1) Les deux remarques de la section 4.1.5.5.B.1 s'appliquent.
- 2) Notre approximation fournit des résultats similaires à ceux de l'approximation de Matthews et Baker dans le cas de petites déformations (voir §4.4.2.1.B-(2i)).

### B.4) Synthèse à propos des paramètres de $\tilde{T}^{gon}$

Les paramètres  $\tilde{\beta}$  de  $\tilde{T}^{gon}$  vérifient (quelle que soit l'approximation) :

$$\begin{cases} \tilde{\beta}_i = \beta_i, & \forall i \in [1, 2N]_{\mathbb{N}} \setminus \{2n-1, 2n\} \\ [\tilde{\beta}_{2n-1} \ \tilde{\beta}_{2n}]^T = S_n'' - S_n \end{cases}.$$

En écrivant :

$$S_n'' - S_n = \underbrace{S_n'' - T(S_n, \beta)}_{\triangle_{\beta^{(n)}}} + \underbrace{T(S_n, \beta) - S_n}_{=[\beta_{2n-1} \ \beta_{2n}]^T},$$

alors :

$$\tilde{\beta} = \beta + \Delta\beta', \quad (4.40)$$

où :

$$\Delta\beta' = [0_{1, 2n-2} \ (\Delta\beta'^{(n)})^T \ 0_{1, 2N-2n}]^T. \quad (4.41)$$

## 4.2 Maillage et suivi : état de l'art

### 4.2.1 Méthodes mono-problème (globales)

Elles consistent à résoudre le problème  $(\mathcal{P}^s)$  (équation (2.52)), que nous rappelons ici :

$$(\mathcal{P}^s) \quad \begin{cases} \min & \frac{1}{2} \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta)) - I_{ref}(\mathbf{m}))^2 \\ \beta \in \mathbb{R}^p \end{cases}, \quad (4.42)$$

par les algorithmes itératifs bien rodés des moindres carrés non linéaires [Bjö96]. Les travaux [MPL00, Mar00, Dud96] s'inscrivent dans ce cadre et, par exemple, Marquant *et al.* [MPL00] implémentent l'algorithme de Gauss-Newton [[DS83], éq. (10.2.3)] (voir §2.4.3.1) avec les spécifications :

- $\Omega$  est l'image toute entière,
- $I_{ref} = I_{t-1}$ ,
- $T$  est une transformation globale du maillage,
- $\beta \in (\mathbb{R}^2)^N$ , où  $N$  est le nombre de nœuds du maillage.

La complexité d'une telle approche est en  $O(mnp^2)$ , où  $m$  et  $n$  désignent les dimensions de l'image et  $p$  le nombre de paramètres à estimer (ici,  $p = 2N$ ).

## 4.2.2 Méthodes pluri-problèmes (locales)

Elles traitent autant de problèmes d'optimisation que de nœuds (ou de mailles) du maillage.

### 4.2.2.1 Algorithme de Nakaya et Harashima

#### A) La théorie

##### A.1) Le principe informel

Nakaya et Harashima [[NH94], §IV-B, page 348] traitent séparément chaque nœud  $S_n$  du maillage  $\mathcal{M}$  en deux temps :

- une étape de « d'approche » (*block matching*), à savoir l'appariement d'un motif  $B_{ref}^{(n)}$  centré en  $S_n$  sur  $I_{ref}$  avec un motif  $B_t^{(n)+d}$ , obtenu par translation globale  $d$ , sur  $I_t$ .
- une étape de « finition » (*hexagonal matching*), à savoir l'appariement du polygone support  $\Pi_n$  du nœud  $S_n$  déplacé de  $d$  sur  $I_{ref}$  avec le polygone support  $T_{\beta(n)}(\Pi_n)$ , où  $\beta(n) = d$ , obtenu par transformation nodale  $T_{\beta(n)}$ , sur  $I_t$ .

##### A.2) Les équations

###### Concernant l'étape « d'approche »

On reprend le problème de l'équation (4.42), pour chaque nœud  $S_n$ , avec :

- $\Omega = B_{ref}^{(n)}$ . Par exemple,  $\Omega = B_{\infty}^{(d)}(S_n, \rho^{(s)})$ , où  $\rho^{(s)} = 1, 2, 3, \dots$  ;
- $\begin{cases} T : \mathbb{R}^2 \times B_{\infty}^{(d)}(O, \rho) & \rightarrow \mathbb{R}^2 \\ (\mathbf{m}, \beta) & \mapsto T(\mathbf{m}, \beta) = \mathbf{m} + \beta \end{cases}$ , où  $B_{\infty}^{(d)}(O, \rho) \subset \mathbb{R}^2$  est habituellement appelée fenêtre de recherche (par exemple,  $\rho = 7$ ) ;
- $I_{ref} = I_{t-1}$ .

###### Concernant l'étape de « finition »

On reprend le problème de l'équation (4.42), pour chaque nœud  $S_n$ , avec :

- $\Omega = \Pi_n$  ;
- $\begin{cases} T : \mathbb{R}^2 \times B_{\infty}^{(d)}(O, \rho) & \rightarrow \mathbb{R}^2 \\ (\mathbf{m}, \beta^{(n)}) & \mapsto T(\mathbf{m}, \beta) = T_{\beta(n)}(\mathbf{m}) \end{cases}$ , où  $B_{\infty}^{(d)}(O, \rho) \subset \mathbb{R}^2$  est similairement appelée fenêtre de recherche (par exemple,  $\rho = 3$ ) et  $T_{\beta(n)}$  est une transformation nodale ;
- $I_{ref} = I_{t-1}$ .

## B) La pratique

i) Illustration : voir figure 4.24.

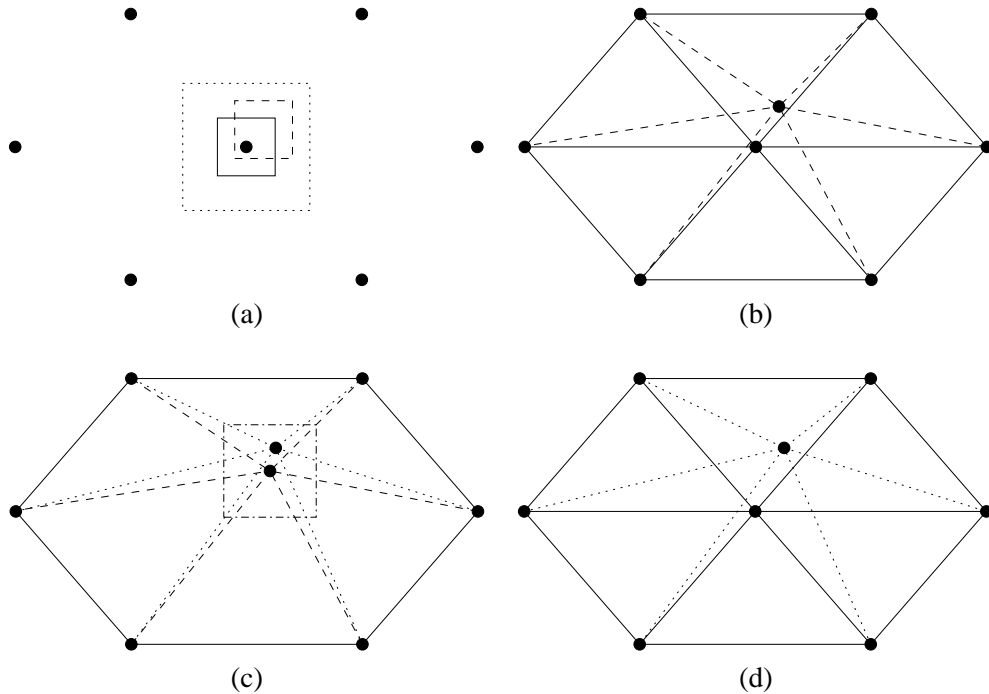


FIG. 4.24 – Les nœuds '•'. (a) Les composants du *block matching* : en traits pleins, le motif à apparier ; en tirets, le motif translaté ; en pointillé, la fenêtre de recherche du motif homologue. (b) Résultat du *block matching* : en traits pleins, le maillage initial  $\mathcal{M}$  ; en traits discontinus, le maillage  $\mathcal{M}_{bm}$  résultant du *block matching*. (c) Les composants de l'*hexagonal matching* : en tirets, le maillage  $\mathcal{M}_{bm}$  ; en pointillé, le maillage résultant de  $T_{\beta(n)}$  ; en point-tiret la fenêtre de recherche de l'*hexagonal matching*. (d) Résultat de l'*hexagonal matching* : en traits pleins, le maillage initial  $\mathcal{M}$  ; en pointillé, le maillage  $\mathcal{M}_{hm}$  résultant de l'*hexagonal matching*.

- 2i) Le traitement des nœuds étant séquentiel, il est nécessaire de le répéter jusqu'à satisfaire un critère de stabilité (quelques itérations). Il serait judicieux de changer l'ordre de parcours des nœuds.
- 3i) À cause de l'exhaustivité des recherches, cette méthode est très coûteuse en temps de calcul [[AT97b], tab. I, page 1265]. Qualitativement parlant, elle se positionne bien [[AT97b], tab. II, page 1268]. La complexité algorithmique est détaillée à la section 4.3.4, où elle est comparée à celle de notre méthode.

### 4.2.2.2 Algorithmes d'Altunbasak et Tekalp

#### A) La théorie

##### A.1) Le principe informel

### A.1.1) À propos des parcours

Étant donné un maillage triangulaire  $\mathcal{M}_{t-1}$  associé à une image  $I_{t-1}$ , Altunbasak et Tekalp [AT97b] distinguent deux types de parcours de  $\mathcal{M}_{t-1}$  pour prédire le maillage  $\mathcal{M}_t$  associé à l'image  $I_t$  (d'une séquence), satisfaisant à :

- i)  $\mathcal{M}_t$  a même topologie que  $\mathcal{M}_{t-1}$  ;
- 2i)  $\mathcal{M}_t = T(\mathcal{M}_{t-1}, \beta)$ , c'est-à-dire est obtenu par transformation affine par morceaux (triangles).

Le premier type de parcours ( $\mathcal{T}$  comme triangle) consiste à visiter séquentiellement, une et une seule fois, dans un ordre prédéfini, chacun des triangles de  $\mathcal{M}_{t-1}$ .

Le second type de parcours ( $\mathcal{N}$  comme nœud) consiste à visiter séquentiellement, une et une seule fois, dans un ordre prédéfini, chacun des nœuds de  $\mathcal{M}_{t-1}$ .

Lors du parcours ( $\mathcal{T}$ ), au cours du traitement d'un triangle  $\Delta$ , tout sommet  $S_n$  de  $\Delta$  appartenant aussi à un  $\Delta_j$  déjà visité induit une contrainte sur la transformation affine  $T_{Aff}$  de  $\Delta$ .

Lors du parcours ( $\mathcal{N}$ ), le traitement d'un nœud  $S_n$  consiste à calculer en parallèle toutes les transformations affines  $(T_{Aff_i})_{i \in J_n}$  associées aux triangles  $\Delta_i$  du polygone support  $\Pi_n$  de  $S_n$  en intégrant les contraintes induites par les nœuds appartenant à plusieurs triangles. Souvenons-nous que  $S_n$  est sommet de  $|J_n|$  triangles !

### A.1.2) À propos des données déduites de $(I_{t-1}, I_t)$

Pour les deux parcours ( $\mathcal{T}$ ) et ( $\mathcal{N}$ ), Altunbasak et Tekalp distinguent à nouveau deux situations.

Le premier cas ( $\mathcal{A}$  comme appariement) est celui où, pour tout triangle  $\Delta_i$ , on dispose de  $n_{\Delta_i} (\geq 3)$  couples de points homologues  $(\mathbf{m}_{ij}, \mathbf{m}'_{ij})_{j=1..n_{\Delta_i}}$ , ce qui permet d'écrire un critère d'inadéquation d'une transformation affine  $T_{Aff_i}$ , à savoir :

$$F_i^{\mathcal{A}}(\beta) = \frac{1}{2} \sum_{j=1}^{n_{\Delta_i}} \|\mathbf{m}'_{ij} - T_{Aff_i}(\mathbf{m}_{ij}, \beta)\|_2^2, \quad (4.43)$$

où  $\beta \in \mathbb{R}^6$ .

Le second cas ( $\mathcal{G}$  comme gradient) est celui où pour tout triangle  $\Delta_i$ , on dispose de  $n_{\Delta_i} (\geq 6)$  triplets  $(\mathbf{m}_{ij}, \nabla I_t(\mathbf{m}_{ij}), \frac{\partial I_t}{\partial t}(\mathbf{m}_{ij}))_{j=1..n_{\Delta_i}}$ , où la dérivée de  $I_t$  par rapport au temps  $\frac{\partial I_t}{\partial t}(\mathbf{m}_{ij})$  peut être approchée par  $I_t(\mathbf{m}_{ij}) - I_{t-1}(\mathbf{m}_{ij})$ , ce qui permet d'écrire un critère d'inadéquation de la contrainte du flot optique [[Hor86], dernière équation, p. 282], à savoir :

$$F_i^{\mathcal{G}}(\beta) = \frac{1}{2} \sum_{j=1}^{n_{\Delta_i}} \left( (\nabla I_t(\mathbf{m}_{ij}))^T T_{Aff_i}(\mathbf{m}_{ij}, \beta) + \frac{\partial I_t}{\partial t}(\mathbf{m}_{ij}) \right)^2, \quad (4.44)$$

où  $\beta \in \mathbb{R}^6$ .

## A.2) Les équations

Par définition de toute transformation affine :

- les deux critères (4.43) et (4.44) sont ceux d'un problème aux moindres carrés linéaires de la forme générique :

$$\begin{cases} \min & \frac{1}{2} \|A\beta - b\|_2^2 \\ \beta \in \mathbb{R}^p \end{cases}.$$

Par exemple, pour le critère (4.43), en notant  $\mathbf{m}_{ij} = [x_{ij} \ y_{ij}]^T$  et de même pour  $\mathbf{m}'_{ij} = [x'_{ij} \ y'_{ij}]^T$ , on a :

$$A = \begin{bmatrix} x_{i1} & y_{i1} & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{in_{\Delta_i}} & y_{in_{\Delta_i}} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{i1} & y_{i1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & x_{in_{\Delta_i}} & y_{in_{\Delta_i}} & 1 \end{bmatrix}, \quad b = \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{in_{\Delta_i}} \\ y'_{i1} \\ \vdots \\ y'_{in_{\Delta_i}} \end{bmatrix} \quad \text{et} \quad \beta = \begin{bmatrix} a_{11} \\ a_{12} \\ b_1 \\ a_{21} \\ a_{22} \\ b_2 \end{bmatrix} \quad (4.45)$$

- toute contrainte exprimant l'appartenance d'un sommet  $S$  à deux triangles  $\Delta_i$  et  $\Delta_j$  est un système linéaire de deux équations à  $2 \times 6$  inconnues.

En conséquence, tout problème de minimisation des critères (4.43) et (4.44) est un problème aux moindres carrés linéaires avec contraintes linéaires d'égalité [[Bjö96], §5.1, p. 187–194].

Il reste à préciser pour chacune des configurations  $(\mathcal{T} + \mathcal{A})$ ,  $(\mathcal{T} + \mathcal{G})$ ,  $(\mathcal{N} + \mathcal{A})$  et  $(\mathcal{N} + \mathcal{G})$  :

- les contraintes,
- l'algorithme de résolution du problème aux moindres carrés linéaires avec contraintes linéaires d'égalité.

### A.2.1) La configuration $(\mathcal{T} + \mathcal{A})$

#### i) Les contraintes

Soit  $\Delta$  le triangle courant du parcours  $(\mathcal{T})$ . Quatre cas de figure peuvent se présenter en fonction du nombre de sommets de  $\Delta$  déjà visités au cours du traitement des triangles déjà parcourus :

- a) zéro sommet : il n'y a aucune contrainte ;
- b) un sommet  $S$  : il y a deux ( $= 1 \times 2$ ) contraintes (car on est dans  $\mathbb{R}^2$ ) traduisant  $S \in \Delta \cap \Delta_j$  ;
- c) deux sommets  $S_1$  et  $S_2$  : il y a quatre ( $= 2 \times 2$ ) contraintes traduisant  $S_1 \in \Delta \cap \Delta_i$  et  $S_2 \in \Delta \cap \Delta_j$  ;
- d) trois sommets  $S_1$ ,  $S_2$  et  $S_3$  : il y a six ( $= 3 \times 2$ ) contraintes traduisant  $S_1 \in \Delta \cap \Delta_i$ ,  $S_2 \in \Delta \cap \Delta_j$  et  $S_3 \in \Delta \cap \Delta_k$ .

#### 2i) L'algorithme de résolution

Altunbasak et Tekalp choisissent la méthode d'élimination analytique des variables dans les contraintes suivie d'une substitution analytique dans le critère débouchant sur une résolution au sens des moindres carrés linéaires d'un problème à, respectivement, 6, 4, 2 inconnues pour les cas (a), (b), (c) et sur un système linéaire de 6 équations à 6 inconnues dans le cas (d).

---

**Remarques**

- 1) L'équation (4.45) montre que le critère est à variables séparables, donc on peut avantageusement remplacer chacun des quatre problèmes à  $2p$  variables par deux problèmes à  $p$  variables (ici,  $p = n_{\Delta_i}$ ).
  - 2) La résolution analytique est un effort « louable » (aux erreurs de calcul près) mais il existe des méthodes numériques très performantes [[Bjö96], §5.1.2, p. 188].
- 

**A.2.2) La configuration ( $\mathcal{T} + \mathcal{G}$ )**
**i) Les contraintes**

Ce sont les mêmes qu'en A.2.1-(i).

**2i) L'algorithme de résolution**

Altunbasak et Tekalp poussent ici les calculs analytiques jusqu'à calculer les équations normales de Gauss pour déboucher sur la résolution d'un système linéaire à respectivement 6 équations à 6 inconnues, 4 équations à 4 inconnues, 2 équations à 2 inconnues pour les cas (a), (b) et (c). Naturellement, le cas (d) est analogue au cas (d) du A.2.1.

---

**Remarques**

- 1) Le critère  $F_i^{\mathcal{G}}$  n'étant pas à variables séparables, la remarque (1) de A.2.1 ne s'applique plus.
  - 2) Les algorithmes de résolution (aux moindres carrés linéaires) de A.2.1-(2i) sont applicables et vice-versa !
- 

**A.2.3) La configuration ( $\mathcal{N} + \mathcal{A}$ )**
**i) Les contraintes**

Soient :

- $S_n$  le nœud courant du parcours ( $\mathcal{N}$ ) ;
- $(\Delta_j)_{j=1..K_n}$  les triangles ordonnés (dans le sens trigonométrique) du polygone support  $\Pi_n$  de  $S_n$ , ayant donc  $S_n$  comme sommet ;
- $S_1, \dots, S_{K_n}$  les sommets ordonnés (dans le sens trigonométrique) de la frontière du polygone  $\Pi_n$ , tels que  $\forall i = 1..K_n, [S_i S_n] = \Delta_i \cap \Delta_{i+1}$ , où  $\Delta_{K_n+1} \stackrel{n}{=} \Delta_1$ .

Alors deux types de contraintes apparaissent :

- l'équation  $T_{Aff_{i_0}}(S_n, \beta_{i_0}) = T_{Aff_j}(S_n, \beta_j), \forall j \in [1, K_n]_{\mathbb{N}} \setminus \{i_0\}$  induit  $2(K_n - 1)$  contraintes affines (car nous sommes dans  $\mathbb{R}^2$ ) ;
- les équations  $T_{Aff_i}(S_k, \beta_i) = T_{Aff_{i+1}}(S_k, \beta_{i+1}), \forall S_k$  tel que

$$[S_n, S_k] = \Delta_i \cap \Delta_{i+1}, \forall i = 1..K_n,$$

où  $\Delta_{K_n+1} \stackrel{n}{=} \Delta_1$  et  $\beta_{K_n+1} \stackrel{n}{=} \beta_1$  induisent  $2K_n$  contraintes affines.

**2i) Le critère**

$$F^{\mathcal{N}+\mathcal{A}}(\beta_1, \dots, \beta_{K_n}) = \sum_{i=1}^{K_n} F_i^{\mathcal{A}}(\beta_i), \quad (4.46)$$

où  $(\beta_1, \dots, \beta_{K_n}) \in (\mathbb{R}^6)^{K_n}$ .

3i) **L'algorithme de résolution**

Comme dans la configuration  $(\mathcal{T} + \mathcal{A})$ , Altunbasak et Tekalp choisissent la méthode d'élimination analytique des variables dans les contraintes suivie d'une substitution analytique dans le critère. Prenant conscience de la propriété de séparation des variables dans le critère, Altunbasak et Tekalp résolvent **deux** problèmes aux moindres carrés linéaires à  $K_n + 1$  ( $= 3K_n - (K_n - 1) - K_n$ ) inconnues.

**A.2.4) La configuration  $(\mathcal{N} + \mathcal{G})$** i) **Les contraintes**

Ce sont les mêmes qu'en A.2.3-(i) ! Altunbasak et Tekalp, qui entrelacent les contraintes du premier type avec celles du second type en parcourant les  $K_n$  arêtes « mitoyennes » de  $\Pi_n$ , en sont-ils conscients ?

2i) **Le critère**

$$F^{\mathcal{N}+\mathcal{G}}(\beta_1, \dots, \beta_{K_n}) = \sum_{i=1}^{K_n} F_i^{\mathcal{G}}(\beta_i), \quad (4.47)$$

où  $(\beta_1, \dots, \beta_{K_n}) \in (\mathbb{R}^6)^{K_n}$ .

3i) **Les algorithmes de résolution**

Le problème de minimisation est donc un problème aux moindres carrés linéaires à  $2(K_n + 1)$  ( $= 6K_n - 2(K_n - 1) - 2K_n$ ) inconnues, si l'on opère la même procédure d'élimination-substitution qu'en A.2.3-(3i).

Contre toute attente et toute recommandation des ouvrages de référence en algorithmique et optimisation numériques, Altunbasak et Tekalp choisissent de voir le problème aux moindres carrés linéaires avec contraintes linéaires d'égalité comme un problème quadratique général avec contraintes affines et de résoudre « frontalement » le système K.K.T standard [[NW06], éq. (16.5), p. 451].

**B) La pratique**

- 1) Il est regrettable qu'aussi bien Altunbasak que Tekalp se soient réfugiés derrière l'argument « 1997, c'est trop loin » pour refuser :
  - de nous communiquer les programmes associés à leur travail,
  - d'apporter des précisions sur les algorithmes implémentés, ne serait-ce que pour savoir s'il s'agit de l'opérateur '\', de la fonction *pinv*, de la fonction *lsqlin...* de MATLAB.
- 2) Pour le parcours  $(\mathcal{T})$ , le résultat global **dépend de l'ordre** de visite des triangles. Altunbasak et Tekalp proposent de choisir en priorité :
  - i) pour  $(\mathcal{T} + \mathcal{A})$ , les triangles les plus fiables au vu des résultats de l'appariement [[AT97b], éq. (19), p. 1260],
  - 2i) pour  $(\mathcal{T} + \mathcal{G})$ , a priori le même ordre que pour  $(\mathcal{T} + \mathcal{A})$ . Mais qu'en est-il lorsqu'on ne dispose pas des appariements ?
- 3) Pour le parcours  $(\mathcal{N})$ , l'ordre de visite des nœuds n'influence, a priori, pas le résultat final. Par contre, un autre effet de bord est inévitable : chaque polygone support

étant traité indépendamment de tous les autres, la topologie n'est plus respectée. Plus exactement, un nœud  $S_n$  peut avoir  $K_n$  images !

Altunbasak et Tekalp proposent une procédure de rétablissement de la topologie qui consiste à conserver comme image de  $S_n$  le barycentre des points  $(T_{Aff_i}(S_n, \beta_i))_{i=1..K_n}$  affectés des poids  $\frac{1}{\|S_n - S_i\|_2}$ , où  $\Delta_i \cap \Delta_{i+1} = [S_n S_i]$ .

Naturellement, au cours de ce post-traitement, l'ordre de visite des nœuds  $S_n$  influence le résultat global et Altunbasak et Tekalp font le même choix qu'en B.2-(i) (en étendant le calcul du critère de fiabilité à tous les triangles  $\Delta_i$  de  $\Pi_n$ ). Là encore, qu'en est-il lorsque l'on ne dispose pas des appariements ?

## 4.3 Maillage et suivi d'objets déformables : notre proposition

### 4.3.1 Introduction

#### 4.3.1.1 Nos objectifs

- suivre dans une séquence d'images  $\{I_0, \dots, I_t, \dots\}$ ,  $t \in \mathbb{N}$ , des **objets déformables** en utilisant un **maillage**, autrement dit, estimer pour chaque image  $I_t$  les paramètres  $\beta_t \in (\mathbb{R}^2)^N$  de la transformation globale  $T$  (équation (4.33)) ;
- conserver la topologie du maillage tout au long du suivi ;
- travailler sur un ensemble de points épars ;
- obtenir un algorithme rapide.

#### 4.3.1.2 Nos hypothèses

Nous supposons les hypothèses suivantes satisfaites :

- les déformations entre deux images consécutives  $I_{t-1}$  et  $I_t$  sont petites ;
- les niveaux de gris restent constants entre  $I_0$  et  $I_t$  aux points d'intérêt (hypothèse temporellement forte et spatialement faible) ;
- le polygone support de chaque nœud (figure 4.25) se déforme globalement de façon affine entre  $I_0$  et  $I_t$ . Nous verrons dans la partie expérimentation que nous pouvons relâcher cette hypothèse.

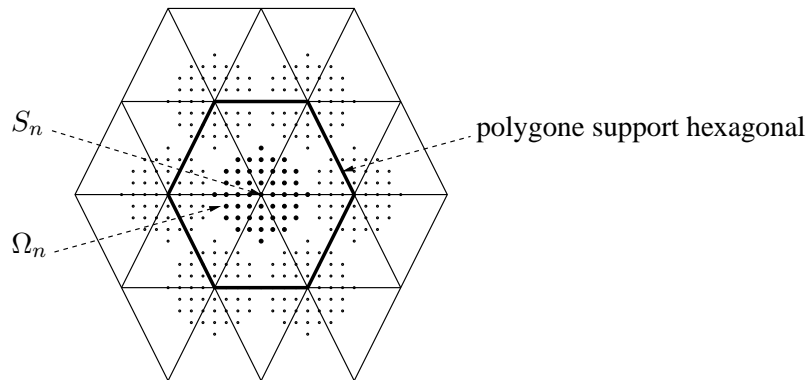


FIG. 4.25 – Exemple de motif d'intérêt  $\Omega_n$ , associé au nœud  $S_n$ .



### 4.3.1.3 Nos leviers

Notre algorithme (CMPT) :

- utilise un maillage triangulaire ;
- combine des méthodes pluri-problèmes (locales) et mono-problème (globale).

(CMPT) peut être vu comme une généralisation de l'algorithme de Jurie et Dhome (JD) [JD02] (voir §2.5.1), plus exactement de notre variante (PP) (voir §2.5.2), au cas du suivi d'objets déformables et se décompose donc en deux étapes :

- une étape d'apprentissage artificiel hors ligne estimant, pour chaque nœud  $S_n$ , un prédicteur linéaire  $A_n$  ;
- une étape de suivi en ligne exploitant astucieusement les prédicteurs  $A_n$ .

## 4.3.2 À propos de l'apprentissage artificiel

### 4.3.2.1 La théorie

#### A) Les principes informels

Nous appliquons notre étape d'apprentissage artificiel (PP\_A) (voir §2.5.2.2) pour chaque **nœud**  $S_n$  ( $n \in [1, N]_{\mathbb{N}}$ ) du maillage  $\mathcal{M}_{ref}$  : il en résulte  $N$  prédicteurs linéaires  $A_n$ , totalement indépendants les uns des autres.

#### B) Les équations

Pour tout nœud  $S_n$ , (PP\_A) calcule le prédicteur  $A_n$  satisfaisant à l'équation générique :

$$\Delta\beta^{(n)} = A_n \Delta \mathbf{I}, \quad (4.48)$$

à partir de :

- $\Delta\beta^{(n),l} \in \mathbb{R}^2, \forall l \in [1, N_a]_{\mathbb{N}}$ , où  $N_a$  est le nombre d'expériences (de perturbations) mises en jeu pour l'apprentissage ;
- 

$$\Delta \mathbf{I}^l = [I_0(T_{\Delta\beta^{(n),l}}^{-1}(\mathbf{m}_j)) - I_0(\mathbf{m}_j)]_{j=1..|\Omega_n|}, \forall l \in [1, N_a]_{\mathbb{N}}. \quad (4.49)$$

- $\Omega_n = \{m_j\}_{j=1..|\Omega_n|}$ , le motif d'intérêt associé à  $S_n$  (voir figure 4.25) ;
- $T$ , une transformation nodale paramétrée par  $\Delta\beta^{(n)}$ .

---

#### Remarques

- 1) Pour simplifier les notations, nous utilisons volontairement  $A_n$  (respectivement  $\Delta \mathbf{I}^l$ ,  $\Delta\beta^{(n)}$ ) en lieu et place de  $A_n^-$  (respectivement  $(\Delta \mathbf{I}^l)^-$ ,  $\Delta\beta^-$ ).
  - 2) Le terme  $T_{\Delta\beta^{(n),l}}^{-1}$  dans l'équation (4.49) justifie la définition, la caractérisation et l'application au maillage de référence  $\mathcal{M}_{ref}$  d'une transformation nodale inverse (voir figure 4.22 et §4.1.5.4).
-

### 4.3.2.2 La pratique

#### A) L'algorithme d'apprentissage : algorithme 4.3.1

Il n'y a rien de plus à dire que de résoudre le problème (2.109) présenté à la section 2.5.1.2 avec les bons composants :

- $B \in \mathbb{R}^{2 \times N_a}$ ,
- $C = [\Delta \mathbf{I}^1 \quad \dots \quad \Delta \mathbf{I}^{N_a}]$ .

Initialisations :

$$C = []$$

$$B = []$$

Algorithme :

**Pour** chaque perturbation  $\Delta \beta^{(n),l} \in \mathbb{R}^2, l \in [1, N_a]_{\mathbb{N}}$  **Faire**

$$\Delta \mathbf{I}^l \leftarrow [I_0(T_{\Delta \beta^{(n),l}}^{-1}(\mathbf{m}_j)) - I_0(\mathbf{m}_j)]_{j=1..|\Omega_n|} \text{ (cf. éq. (4.49))}$$

$$C \leftarrow \begin{bmatrix} C & \Delta \mathbf{I}^l \end{bmatrix}$$

$$B \leftarrow \begin{bmatrix} B & \Delta \beta^{(n),l} \end{bmatrix}$$

**Fin Pour**

En notant  $B = [b_1 \quad b_2]^T$ , résoudre le problème aux moindres carrés linéaires :

$$A_n \leftarrow \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \text{ où } \forall i = 1, 2, a_i = \begin{cases} \operatorname{argmin} & \|C^T a - b_i^T\|^2 \\ a \in \mathbb{R}^{|\Omega_n|} \end{cases}$$

Algorithme 4.3.1: Algorithme d'apprentissage artificiel : estimation du prédicteur  $A_n$ .

#### B) Les détails de l'implémentation

Dans nos expériences,  $N_a$  est de l'ordre de 2500 et l'amplitude maximale des mouvements de 10 pixels ( $\|\Delta \beta^{(n),l}\|_{\infty} \leq 10$ ).

### 4.3.3 À propos du suivi « au fil de l'eau »

#### 4.3.3.1 La théorie

##### A) Les principes informels

L'estimation de la déformation globale du maillage passe par l'application de l'étape de suivi (PP\_S) (voir §2.5.2.3), indépendamment pour chaque nœud  $S_n$ , afin d'obtenir la prédiction de son déplacement hypothétique  $\Delta \beta'^{(n)}$ . Puis le nœud minimisant le critère fondamental du suivi ( $\mathcal{P}^s$ ) est sélectionné et déplacé. Ce processus est itéré via un **algorithme de relaxation** jusqu'à convergence.

##### B) Les équations

Notre modèle mathématique associé au problème global s'identifie au problème fondamental du suivi ( $\mathcal{P}^s$ ) (équation (2.52)) que nous rappelons :

$$(\mathcal{P}^s) \quad \begin{cases} \min \\ \beta \in (\mathbb{R}^2)^N \end{cases} \frac{1}{2} F_t(\beta), \quad (4.50)$$

où :

-

$$\begin{cases} F_t : (\mathbb{R}^2)^N \rightarrow \mathbb{R} \\ \beta \mapsto F_t(\beta) = \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta)) - I_0(T(\mathbf{m}, \beta_0)))^2 \end{cases} \quad (4.51)$$

- $\Omega$  est un ensemble de points éparés du domaine du maillage formant le motif d'intérêt (voir figure 4.25) qui peut être le résultat d'une étape de sélection (voir chapitre 3).

Nous décomposons le problème  $(\mathcal{P}^s)$  en sous-problèmes « plus simples » ayant une portée locale, en :

- opérant une partition de  $\Omega$ , à savoir :

$$\begin{cases} \Omega = \bigcup_{i=1}^N \Omega_i \\ \Omega_i \cap \Omega_j = \emptyset, \forall i \neq j \end{cases} \quad (4.52)$$

où  $\forall n \in [1, N]_{\mathbb{N}}$ ,  $\Omega_n$  est le motif d'intérêt associé au nœud  $S_n$  (voir figure 4.25) ;

- associant à chaque nœud  $S_n$  un critère local  $F_t^n$  :

$$\begin{cases} F_t^n : (\mathbb{R}^2)^N \rightarrow \mathbb{R} \\ \beta \mapsto F_t^n(\beta) = \sum_{\mathbf{m} \in \Omega_n} (I_t(T(\mathbf{m}, \beta)) - I_0(T(\mathbf{m}, \beta_0)))^2 \end{cases} \quad (4.53)$$

On a donc :

$$\forall \beta \in (\mathbb{R}^2)^N, F_t(\beta) = \sum_{n=1}^N F_t^n(\beta). \quad (4.54)$$

Nous choisissons de résoudre le problème global via une méthode itérative telle que :

$$\begin{cases} \beta_t^{(0)} = \beta_{t-1} \\ \beta_t^{(k)} = \beta_t^{(k-1)} + \Delta\beta', \quad \forall k \geq 1 \end{cases} \quad (4.55)$$

où  $\Delta\beta'$  est obtenu en résolvant les problèmes locaux, donc  $\Delta\beta'$  est de la forme :

$$\Delta\beta' = [0_{1,2n-2} \ (\Delta\beta'^{(n)})^T \ 0_{1,2N-2n}]^T,$$

où  $S_n$  est le nœud conduisant, à l'itération  $k$ , au minimum du critère global  $F_t$ .

Pour résoudre les problèmes locaux, nous utilisons la méthode de résolution (PP) (voir §2.5.2), donc, à l'itération  $k$ , le critère  $F_t^n$  devient  $F_t^{n,i(nverse)}$  :

$$\begin{cases} F_t^{n,i} : (\mathbb{R}^2)^N \rightarrow \mathbb{R} \\ \Delta\beta \mapsto F_t^{n,i}(\Delta\beta) = \sum_{\mathbf{m} \in \Omega_n} (I_t(T(\mathbf{m}, \beta_t^{(k-1)})) - I_0(T(\mathbf{m}, \beta_0 + \Delta\beta)))^2 \end{cases} \quad (4.56)$$

où  $\Delta\beta = [0_{1,2n-2} \ (\Delta\beta^{(n)})^T \ 0_{1,2N-2n}]^T$ .

La résolution de ce problème avec (PP\_S) donne :

$$T^{\beta_t^{(k)}} \leftarrow T^{\beta_t^{(k-1)}} \circ T^{\Delta\beta}, \quad (4.57)$$

où  $T^{\Delta\beta}$  est une transformation nodale et  $\Delta\beta^{(n)} = A_n \Delta\mathbf{I}$  avec :

$$\Delta\mathbf{I} = [I_t(T(\mathbf{m}_j, \beta_t^{(k-1)})) - I_0(T(\mathbf{m}_j, \beta_0))]_{j=1..|\Omega_n|} \quad (4.58)$$

En utilisant notre approximation de la composition d'une transformation globale et d'une transformation nodale, faisant intervenir les coordonnées barycentriques généralisées (voir §4.1.5.5.B.3), la transformation approchée  $\tilde{T}^{\beta_t^{(k)}}$  de  $T^{\beta_t^{(k)}}$  est paramétrée par :

$$\beta_t^{(k)} = \beta_t^{(k-1)} + \Delta\beta', \quad (4.59)$$

où  $\Delta\beta'$  est issu de l'approximation.

---

#### Remarque

Notons que, dans le cas où  $T$  est une transformation (nodale, globale) d'un maillage, l'approche (PP) est judicieuse puisqu'elle en conserve la topologie tout au long du suivi, ce qui ne serait pas le cas si la transformation inverse était calculée pendant celui-ci.

---

### 4.3.3.2 La pratique

#### A) L'algorithme de suivi : algorithme 4.3.2

Rappelons que les nœuds sont visités dans l'ordre qui privilégie la décroissance maximale du critère global (4.51).

**Initialisations :**

$\Delta\beta^s \leftarrow [0 \dots 0]^T \in (\mathbb{R}^2)^N$  - - pour sauvegarder tous les  $\Delta\beta'^{(n)}$   
 $\widehat{F}_t \leftarrow [0 \dots 0]^T \in \mathbb{R}^N$  - - pour sauvegarder tous les  $F_t$  en supposant que le  
 - - nœud  $S_n$  est déplacé selon  $\Delta\beta'^{(n)}$

- -  $k \leftarrow 0$

$\beta_t \leftarrow \beta_{t-1}$

$F_t \leftarrow \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta_t)) - I_0(T(\mathbf{m}, \beta_0)))^2$  (voir éq. (4.51))

**Algorithme :**

- - À la première itération, il faut prédire le déplacement de tous les nœuds

**Pour** chaque nœud  $S_n, n \in [1, N]_{\mathbb{N}}$  **Faire**

$\Delta\mathbf{I}^{(n)} \leftarrow [I_t(T(\mathbf{m}_i, \beta_t)) - I_0(T(\mathbf{m}_i, \beta_0))]_{i=1..|\Omega_n|}$  (voir éq. (4.58))

$\Delta\beta^{(n)} \leftarrow A_n \Delta\mathbf{I}^{(n)}$

$S''_n \leftarrow \text{approximation\_avec\_cbg}(\mathcal{M}_{ref}, n, \Delta\beta^{(n)}, \beta_t)$

$[\Delta\beta_{2n-1}^s \Delta\beta_{2n}^s]^T \leftarrow S''_n - T(S_n, \beta_t)$

$\beta_{tn} \leftarrow \beta_t + [0_{2n-2,1} \Delta\beta_{2n-1}^s \Delta\beta_{2n}^s 0_{2N-2n,1}]^T$

$\widehat{F}_t(n) \leftarrow \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta_{tn})) - I_0(T(\mathbf{m}, \beta_0)))^2$  (voir éq. (4.51))

**Fin Pour**

- - Sélection du nœud

$n \leftarrow \begin{cases} \text{argmax} & (F_t - \widehat{F}_t(n)) \\ n \in [1, N]_{\mathbb{N}} \end{cases}$

**Tant que**  $(F_t - \widehat{F}_t(n)) > \tau_F$  **Faire**

- - Déplacement du nœud sélectionné

- -  $k \leftarrow k + 1$

$F_t \leftarrow \widehat{F}_t(n)$

$\beta_t \leftarrow \beta_t + [0_{2n-2,1} \Delta\beta_{2n-1}^s \Delta\beta_{2n}^s 0_{2N-2n,1}]^T$

- - Pour toute itération  $k \geq 1$ , il suffit de prédire le déplacement du nœud  $n$  et

- - ceux de son polygone support : le contexte de tous les autres restant constant

$J \leftarrow \text{polygone\_support\_vers\_sommets}(\mathcal{M}_{ref}, n)$

**Pour** chaque nœud  $S_j, j \in \{n\} \cup J$  **Faire**

$\Delta\mathbf{I}^{(j)} \leftarrow [I_t(T(\mathbf{m}_i, \beta_t)) - I_0(T(\mathbf{m}_i, \beta_0))]_{i=1..|\Omega_j|}$  (voir éq. (4.58))

$\Delta\beta^{(j)} \leftarrow A_j \Delta\mathbf{I}^{(j)}$

$S''_j \leftarrow \text{approximation\_avec\_cbg}(\mathcal{M}_{ref}, j, \Delta\beta^{(j)}, \beta_t)$

$[\Delta\beta_{2j-1}^s \Delta\beta_{2j}^s]^T \leftarrow S''_j - T(S_j, \beta_t)$

$\beta_{tj} \leftarrow \beta_t + [0_{2j-2,1} \Delta\beta_{2j-1}^s \Delta\beta_{2j}^s 0_{2N-2j,1}]^T$

$\widehat{F}_t(j) \leftarrow \sum_{\mathbf{m} \in \Omega} (I_t(T(\mathbf{m}, \beta_{tj})) - I_0(T(\mathbf{m}, \beta_0)))^2$  (voir éq. (4.51))

**Fin Pour**

- - Sélection du nœud

$n \leftarrow \begin{cases} \text{argmax} & (F_t - \widehat{F}_t(n)) \\ n \in [1, N]_{\mathbb{N}} \end{cases}$

**Fin Tant que**

**Fonctions utiles :**

$S''_n = \text{approximation\_avec\_cbg}(\mathcal{M}, n, \beta^{(n)}, \beta)$  est détaillé dans l'algorithme 4.1.5.

$J = \text{polygone\_support\_vers\_sommets}(\mathcal{M}, n)$  : renvoie les indices des nœuds de  $\mathcal{M}$ , ordonnés selon le sens trigonométrique, qui sont sommets de  $\Pi_n$  (même fonction que dans l'algorithme 4.1.5).

Algorithme 4.3.2: Algorithme de suivi.

## B) Les détails de l'implémentation

- 1) Il est possible d'éviter l'évaluation d'expressions constantes dans le corps des deux schémas itératifs « **Pour** », en remplaçant :

- dans l'étape d'initialisations,  $F_t$  par  $[F_{t|\Pi_1 \cap \Omega} \cdots F_{t|\Pi_N \cap \Omega}]^T$ , où :

$$F_{t|\Pi_n \cap \Omega} = \frac{1}{|\Pi_n \cap \Omega|} \sum_{\mathbf{m} \in \Pi_n \cap \Omega} (I_t(T(\mathbf{m}, \beta_t)) - I_0(T(\mathbf{m}, \beta_0)))^2 \quad (4.60)$$

-  $\Omega$  par  $\Pi_k \cap \Omega$  dans l'expression normalisée de  $\widehat{F}_t(k)$ , à savoir :

$$\widehat{F}_t(k) = \frac{1}{|\Pi_k \cap \Omega|} \sum_{\mathbf{m} \in \Pi_k \cap \Omega} (I_t(T(\mathbf{m}, \beta_{tk})) - I_0(T(\mathbf{m}, \beta_0)))^2 \quad (4.61)$$

-  $F_t$  par  $F_t(n)$ .

- 2) Typiquement,  $\tau_F = 10^{-7}$ .

### 4.3.4 À propos de la complexité

En terme de complexité algorithmique, comparée à l'algorithme (NH) de Nakaya et Harashima (voir §4.2.2), notre méthode (CMPT) est moins complexe. En effet, alors que pour l'algorithme NH les recherches exhaustives sont à effectuer entre chacune des images successives, pour notre méthode, l'apprentissage est fait une seule fois pour toute la séquence d'images puis, pour le suivi, il ne reste plus qu'à calculer des produits « matrice-vecteur » et à utiliser les coordonnées barycentriques généralisées pour chaque image.

En effet, pour tous les nœuds, l'algorithme NH utilise de l'ordre de :

$$NM_{bm}B + 3L_{hm}M_{hm}P_{hm} \quad (4.62)$$

opérations flottantes par image, où :

- $N$  est le nombre de nœuds du maillage,
- $M_{bm}$  est le nombre de déplacements testés pour l'appariement de blocs lors de l'étape « d'approche »,
- $B = \rho$  est la dimension du bloc de corrélation pour l'appariement de blocs,
- $L_{hm}$  est le nombre d'itérations lors de l'étape « de finition »,
- $M_{hm}$  est le nombre de déplacements testés lors de l'étape « de finition »,
- $P_{hm}$  est le nombre de pixels du domaine du maillage  $P_{hm} = |\Omega|$ ,

et, pour notre méthode, le nombre d'opérations flottantes est, pour l'étape d'apprentissage (effectuée une seule fois), de l'ordre de :

$$N_a^2 P, \quad (4.63)$$

où  $N_a$  est le nombre de perturbations, et ensuite, pour l'étape de suivi (par image), de l'ordre de :

$$(4 + \frac{3LK}{N})P, \quad (4.64)$$

où :

- $L$  est le nombre d'itérations,
- $K$  est le nombre de triangles du polygone support d'un nœud,

- $P$  est le nombre de points d'intérêt du maillage  $P = \sum_{i=1}^N |\Omega_n|$ .

Pour l'algorithme NH,  $P_{hm}$  représente en règle général l'ensemble des pixels du maillage. En supposant que  $M_{bm} = M_{hm}$ , et étant donné que  $L_{hm}$  est généralement un multiple de  $N$  car tous les nœuds sont étudiés successivement plusieurs fois, nous pouvons négliger la partie sur l'appariement de blocs. Le nombre d'opérations flottantes de l'algorithme NH est donc de l'ordre de :

$$3L_{hm}M_{hm}P_{hm}. \quad (4.65)$$

Pour notre méthode,  $P$  est un ensemble éparé de points. L'étape d'apprentissage a pour complexité celle du calcul de la pseudo-inverse. En ce qui concerne l'étape de suivi,  $K = 6$  (dans notre cas) et suivant les images, nous pouvons aussi bien avoir  $L/N > 1$  que  $L/N < 1$ . Nous ne pouvons donc pas trop simplifier ce nombre d'opérations flottantes. Mais en gros, le nombre d'opérations flottantes du suivi est de l'ordre de :

$$\alpha P \quad (4.66)$$

avec  $\alpha \ll 3L_{hm}M_{hm}$  et de même  $P \leq P_{hm}$ .

Pour donner quelques valeurs numériques, si l'on s'intéresse au cas du suivi de la séquence d'images Miss America présentée dans la partie résultat (voir §4.4) avec un maillage de 19 nœuds et pour l'algorithme NH avec une précision de 0.25 pixel, on a :

- $L_{hm} = 57 (= 3 \times 19)$ ,
- $M_{hm} = 17$ ,
- $P_{hm} = 30201$ ,
- $L = 29.5$  (ceci est une valeur moyenne du nombre d'itérations nécessaires tout au long du suivi sur les 109 images de la séquence),
- $N_a = 2500$ ,
- $P = 23713$  (ceci correspond au cas où tous les pixels du disque autour du nœud composent le motif d'intérêt. Cette valeur décroît rapidement si l'on ne sélectionne qu'un pixel sur 2 ou plus. Par exemple, pour un pixel sur 2, on a  $P = 1468$ ).

On obtient donc  $\alpha \approx 32$  et  $3L_{hm}M_{hm} = 2907$  et on a bien  $23713 < 30201$ . D'où un nombre d'opérations flottantes pour l'algorithme NH d'environ  $8.8 \cdot 10^7$  à chaque image à comparer au nombre d'opérations flottantes pendant l'étape de suivi de l'algorithme (CMPT) d'environ  $7.6 \cdot 10^5$ . Le nombre d'opérations flottantes de l'étape d'apprentissage s'élève à  $1.5 \cdot 10^{11}$ . Il est rentabilisé à partir du moment où l'on travaille sur une séquence de plus de 1000 images pour cet exemple.

## 4.4 Expérimentation

### 4.4.1 Protocole

Pour valider notre méthode de suivi présentée à la section 4.3, nous établissons :

- une **validation locale sur des séquences d'images artificielles** déformées par des transformations affines globales ou par morceaux (les morceaux sont les triangles du maillage). Elle permet en particulier de valider nos choix et de relâcher l'hypothèse forte de transformation affine des polygones supports pour l'utilisation des coordonnées barycentriques généralisées de Wachspress lors du suivi.
- une **validation globale sur des séquences d'images artificielles** générées à partir de transformations non rigides fondées sur des fonctions de base radiales de type « plaque mince » (voir §2.3.1.2).

- une validation globale sur des séquences d'images réelles.

#### 4.4.1.1 Les caractéristiques communes

**Les images de référence.** Elles sont respectivement la première image de deux séquences d'images réelles « Miss America » (MA) et « Foreman » (F). Il s'agit donc respectivement de  $I_0^{(MA)}$  (de dimensions  $288 \times 352$ ) et  $I_0^{(F)}$  (de dimensions  $288 \times 352$ ). Elles sont présentées sur la figure 4.26.

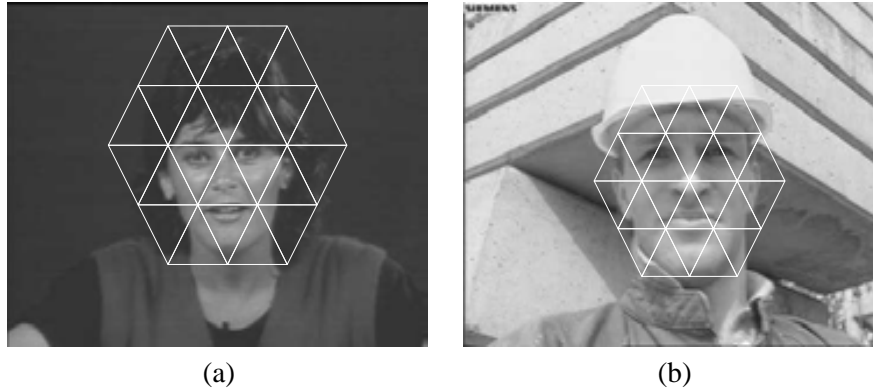


FIG. 4.26 – Les images de référence et leur maillage initial : (a) « Miss America » (MA) ; (b) « Foreman » (F).

**Les maillages initiaux.** La figure 4.26 montre les deux images de référence avec leur maillage initial associé : un maillage régulier de triangles de base et de hauteur 50 pixels pour MA (respectivement 40 pixels pour F).

**Le motif d'intérêt.** Il est formé de points régulièrement répartis autour des nœuds du maillage, ici, tous les pixels du disque de centre le nœud et de rayon 25 pixels pour MA, soit environ 1960 points pour un nœud interne au maillage (respectivement 20 pixels pour F, soit environ 1256 points). Pour chaque image de référence, le motif d'intérêt est illustré sur la figure 4.27.

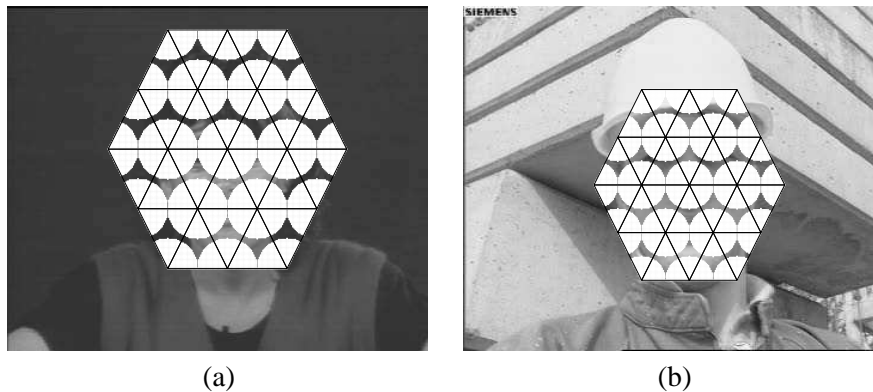


FIG. 4.27 – Les motifs d'intérêt : (a) « Miss America » (MA) ; (b) « Foreman » (F).



**Les perturbations d'apprentissage.** Elles sont les mêmes pour tous les nœuds et les deux images ( $N_a = 2500$ ). Elles suivent une loi uniforme sur  $[-10, 10]$  selon  $x$  et  $y$ .

**Le critère d'arrêt.** Pour l'algorithme de relaxation dans l'étape de suivi (voir algorithme 4.3.2), le seuil  $\tau_F$  du critère d'arrêt est positionné à  $10^{-7}$ .

#### 4.4.1.2 Les séquences d'images

##### A) Les séquences d'images artificielles

Elles sont construites sur le schéma suivant :

- i)  $I_0^{(MA)}$  (respectivement  $I_0^{(F)}$ ) est la première image de la séquence MA (respectivement F) ;
- ii)  $I_t^{(MA)/(F)}$ ,  $t > 0$  est déduite de  $I_{t-1}$  par application d'une transformation  $T$ .

Pour la **validation locale**, la transformation  $T$  peut être :

- une transformation nodale,
- une transformation globale (translation, rotation de centre le milieu de l'image, zoom de point fixe le milieu de l'image, transformation affine, composition de telles transformations),
- une transformation affine par morceaux paramétrée par le déplacement des nœuds du maillage, les morceaux étant les triangles du maillage,
- une composition de telles transformations.

Ces séquences d'images sont nommées « tnX » ou « tgYX », où :

- « tg » (respectivement « tn ») signifie « transformation globale » (respectivement « transformation nodale ») ;
- X désigne le numéro de la séquence ;
- Y désigne le type de transformation qui peut composer la transformation globale  $T$  : translation ( $Y = \text{« t »}$ ), rotation ( $Y = \text{« r »}$ ), zoom ( $Y = \text{« z »}$ ), transformation nodale ( $Y = \text{« n »}$ ) ou plusieurs transformations, par exemple translation et rotation ( $Y = \text{« tr »}$ ).

Par exemple, la séquence :

- « tgt1 » est la séquence d'images artificielles numéro 1 pour laquelle la transformation  $T$  est une translation,
- « tgr3 » est la séquence d'images artificielles numéro 3 pour laquelle la transformation  $T$  est une rotation,
- « tgrz1 » est la séquence d'images artificielles numéro 1 pour laquelle la transformation  $T$  est une rotation, un zoom ou la composition d'une rotation et d'un zoom,
- « tn2 » est la séquence d'images artificielles numéro 2 pour laquelle la transformation  $T$  est une ou plusieurs transformations nodales.

Pour la **validation globale**, la transformation  $T$  est une transformation non rigide fondée sur des fonctions de base radiales de type « plaque mince » et de centres d'influence les nœuds du maillage (voir §2.3.1.2) [[Boo89], §I-D, page 570]. Une telle séquence, sur laquelle nous présentons des résultats, de 55 images de dimensions  $368 \times 432$  est illustrée sur la figure 4.28 pour laquelle l'image de référence est  $I_0^{(F)}$ . On nomme cette séquence

« foreman\_artificiel ». Il est à noter que l'image de référence a été agrandie (voir les bandes noires de 80 pixels à droite et en dessous) afin que le maillage ne sorte pas de l'image lors des transformations. La position du maillage initial est donc bien la même. De même, on étudie une autre séquence d'images artificielles, que l'on nomme « missa\_artificiel », générée à partir de  $I_0^{(MA)}$ .

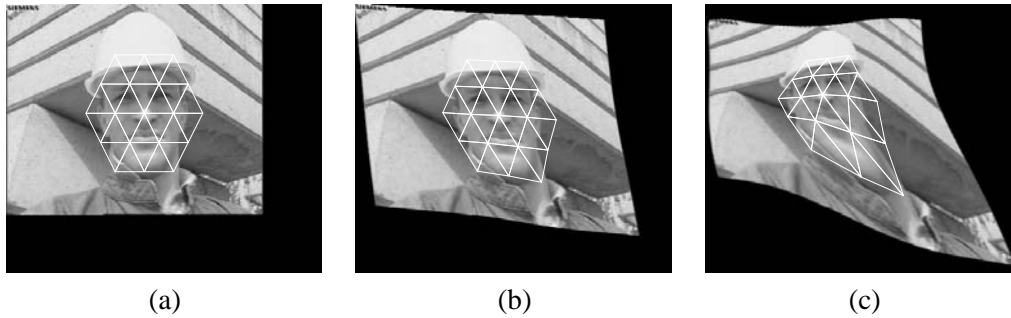


FIG. 4.28 – Quelques images de la séquence d'images artificielles « foreman\_artificiel ». (a) image de référence  $I_0^{(F)}$  et maillage d'apprentissage. (b) image 13 et la vérité terrain du maillage. (c) image 43 et la vérité terrain du maillage.

## B) Les séquences d'images réelles

Nous avons testé notre algorithme de suivi sur les séquences d'images réelles « Miss America » (MA) et « Foreman » (F) au format CIF (image de dimensions  $288 \times 352$ ). Elles contiennent respectivement 109 images pour MA et 250 pour F.

### 4.4.1.3 Les algorithmes de suivi comparés

Les quatre approches d'Altunbasak et Tekalp [AT97b] évoquées à la section 4.2.2 n'ont malheureusement pas pu être comparées à la nôtre. En effet, nous n'avons pas pu les implémenter correctement (des imprécisions des articles restent à être élucidées) et nous regrettons que les auteurs n'aient pas pu nous aider dans cette tâche ou nous fournir leur propre implémentation.

L'algorithme (NH) de Nakaya et Harashima [NH94] est comparé à notre méthode. Nous rappelons que l'algorithme NH donne des résultats meilleurs à ceux d'Altunbasak et Tekalp [[AT97b], tab. II, page 1268]. Il s'agit de notre propre implémentation de leur algorithme pour laquelle :

- l'étape d'approche a été effectuée avec un bloc de dimension  $14 \times 14$  pixels, une fenêtre de recherche de dimension  $15 \times 15$  pixels et un pas de recherche d'un pixel ;
- l'étape de finition a été effectuée avec une fenêtre de recherche de dimension  $5 \times 5$  pixels et un pas de 0.1 pixel pour la séquence de synthèse (respectivement  $1/4$  de pixel pour les séquences réelles ; les temps de calcul étant trop importants, nous n'avons pas fait tourner le suivi avec un pas de 0.1 pixel). Cette étape est itérée trois fois sur chacun des nœuds du maillage. Il est à noter que les nœuds sont traités dans l'ordre de leur numérotation. Ils n'ont pas été ordonnés de façon particulière.

Pour notre algorithme de suivi (CMPT) :

- pour l'étape d'apprentissage, nous étudions et comparons deux méthodes d'apprentissage. La première consiste à appliquer la transformation nodale inverse

exacte ( $\forall \mathbf{m}, \forall \beta^{(n)}, T_{\beta^{(n)}}^{-1}(\mathbf{m})$ ). Nous nommons cet apprentissage : « apprentissage exact » (aE). La deuxième consiste à approcher la transformation nodale inverse par l'opposé du déplacement du noeud, soit  $T_{\beta^{(n)}}^{-1}(\mathbf{m}) \approx T_{-\beta^{(n)}}(\mathbf{m})$ . Notre intérêt pour cette proposition est tout d'abord historique vis-à-vis du déroulement de la thèse mais aussi pragmatique puisqu'elle engendre ainsi un maillage conforme dont les arêtes ne sont pas brisées. Cette deuxième approche permet ainsi de mesurer l'intérêt d'une inversion exacte qui n'avait pas été proposée et donc utilisée avant en vision. Nous nommons cette dernière « apprentissage naïf » (aN) ;

- pour l'étape de suivi, nous étudions et comparons trois approches pour l'approximation de la composition d'une transformation globale du maillage avec une transformation nodale. La première consiste à prendre comme déplacement du noeud celui fourni par le prédicteur linéaire. Nous notons cette approche (sPR) (PR pour prédicteur). La seconde consiste à prendre comme déplacement l'approximation proposée par Matthews et Baker [[MB04], §4.1.4, p. 147] (voir l'équation (4.39)). Nous notons cette approche (sMB). Enfin, la troisième, celle que nous proposons, consiste à utiliser les coordonnées barycentriques généralisées. Nous notons cette approche (sCBG).

Pour faire référence à ces différentes méthodes, nous les nommons selon le principe suivant : CMPT\_a(N | E)\_s(PR | MB | CBG). Ainsi, par exemple, CMPT\_aN\_sMB désigne notre méthode avec un apprentissage naïf et l'utilisation de l'approximation de Matthews et Baker lors du suivi.

#### 4.4.1.4 Validation

Pour valider ces différentes méthodes, nous définissons :

- lorsque la vérité terrain est connue, un critère fondé sur l'écart entre la position des noeuds de la vérité terrain et celle estimée,
- lorsque l'on travaille sur des séquences d'images réelles et donc que rien n'est connu, le critère de qualité de rendu visuel, nommé PRSB (ou PSNR), largement utilisé dans le monde de la compression d'images.

##### A) Critère sur la position des noeuds

Pour les séquences d'images artificielles, la vérité terrain est connue, c'est-à-dire que pour toute image  $I_t (\stackrel{n}{=} f)$ , les paramètres de la transformation du maillage :

$$\beta^{(f)} = [\beta_1^{(f)^T} \dots \beta_N^{(f)^T}]^T \in (\mathbb{R}^2)^N,$$

où :

$$\forall n \in [1, N]_{\mathbb{N}}, \beta_n^{(f)} \in \mathbb{R}^2$$

sont connus.

Dans le cas de la **validation locale**, on suppose que, sur l'image  $I_{t-1} (\stackrel{n}{=} f - 1)$ , le maillage est parfaitement positionné et que sur l'image  $f$ , tous les noeuds du maillage sont parfaitement positionnés **sauf un**, le noeud  $S_n$ . L'estimation des déplacements ne s'effectue donc que sur le noeud  $S_n$ . Le critère de validation  $D_n^{(f)}$ , exprimé en pixels, est donc :

$$D_n^{(f)} = \| \beta_n^{(f)} - \widehat{\beta_n^{(f)}} \|_2, \quad (4.67)$$

où :

- $\beta_n^{(f)} \in \mathbb{R}^2$  est le déplacement du nœud  $S_n$  sur l'image  $f$  fourni par la vérité terrain,
- $\widehat{\beta}_n^{(f)} \in \mathbb{R}^2$  est le déplacement du nœud  $S_n$  sur l'image  $f$  déterminé par un algorithme de suivi.

Dans le cas de la **validation globale**, on suppose que les nœuds du bord  $\mathcal{N}_b$  du maillage sont parfaitement positionnés sur l'image  $f$ , donc l'estimation des déplacements ne s'effectue que sur les nœuds internes  $\mathcal{N}_i$  du maillage. Le critère de validation  $D_{moy}^{(f)}$ , exprimé en pixels, est donc :

$$D_{moy}^{(f)} = \frac{1}{|\mathcal{N}_i|} \sum_{i \in \mathcal{N}_i} \|\beta_i^{(f)} - \widehat{\beta}_i^{(f)}\|_2, \quad (4.68)$$

## B) Critère de validation PRSB (PSNR)

Dans le cas des séquences d'images réelles, la vérité terrain n'étant pas connue, le critère de validation utilisé est une mesure de qualité visuelle d'une image. Il s'agit du pic du rapport signal sur bruit (PRSB ou PSNR, en anglais) exprimé en décibels (en dB). Il est à souligner que la mesure du PRSB est telle que plus le PRSB est grand, meilleur est le rendu visuel. Elle est donnée par :

$$PRSB = -10 \log_{10} \left( \frac{EQM}{(2^b - 1)^2} \right), \quad (4.69)$$

où :

- $b$  représente le nombre de bits sur lequel est codé un pixel. Ici, pour un niveau de gris,  $b$  vaut 8,
- $EQM$  représente l'erreur quadratique moyenne entre l'image  $I_t$  de la séquence et l'image reconstruite  $\hat{I}_t$  à partir de l'image de référence  $I_0$  et du maillage  $\mathcal{M}_t$ , c'est-à-dire le maillage initial  $\mathcal{M}_0$  recalé sur  $I_t$ . On a donc :

$$EQM = \frac{1}{nbl * nbc} \sum_{i=1}^{nbl} \sum_{j=1}^{nbc} (I_t(i, j) - \hat{I}_t(i, j))^2, \quad (4.70)$$

où  $nbl$  (respectivement  $nbc$ ) est le nombre de lignes (respectivement de colonnes) de l'image.

L'image  $\hat{I}_t$  reconstruite est ainsi déterminée :

$$\begin{cases} \forall \mathbf{m} \notin \mathcal{M}_t, & \hat{I}_t(\mathbf{m}) \leftarrow I_t(\mathbf{m}) \\ \forall \mathbf{m} \in \mathcal{M}_t, & \hat{I}_t(\mathbf{m}) \leftarrow I_0(T^{-1}(\mathbf{m}, \beta_t)) \end{cases}. \quad (4.71)$$

Les niveaux de gris aux points de coordonnées non entières sont calculés par interpolation bilinéaire à partir des quatre pixels voisins (cf. remarque de la section 4.1.3.1.2),

## 4.4.2 Résultats

### 4.4.2.1 Validation locale sur des séquences d'images artificielles

#### A) Des chiffres, des courbes

La figure 4.29 montre de façon synthétique une petite partie des résultats obtenus.

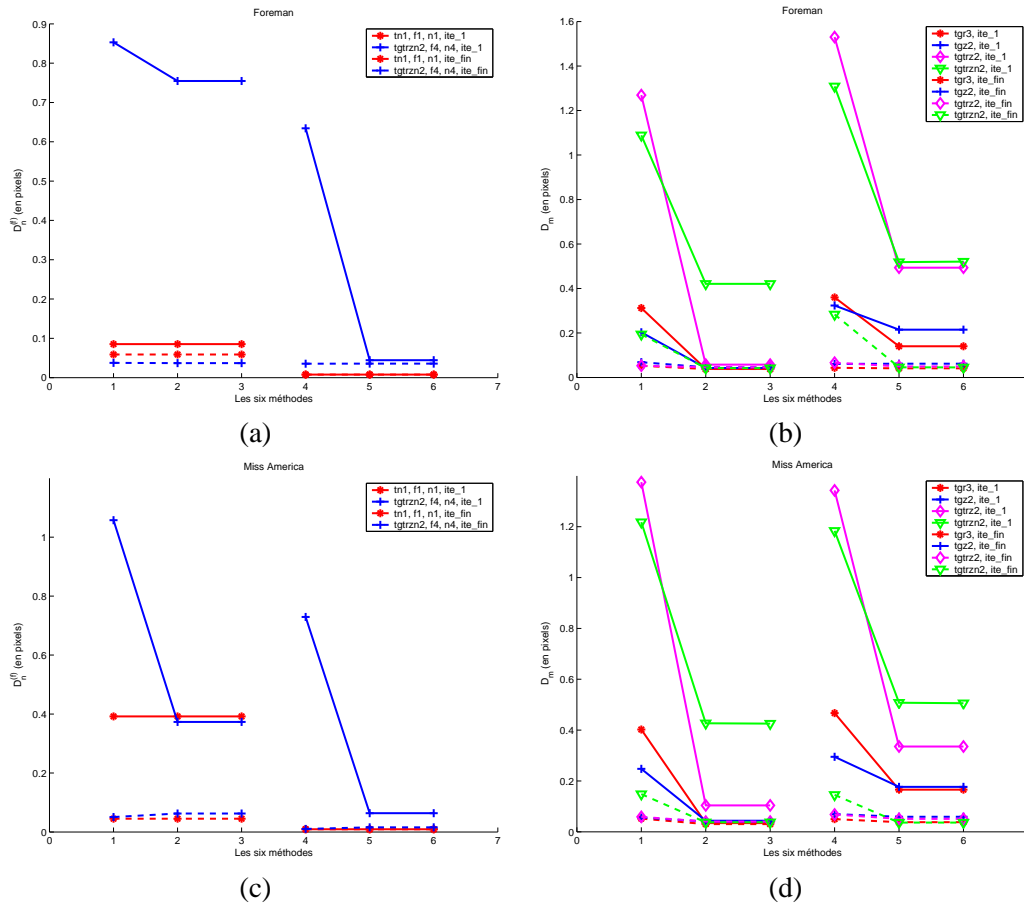


FIG. 4.29 – Validation locale entre six approches de notre suivi sur des séquences d’images artificielles générées à partir de  $I_0^{(F)}$  pour (a) et (b) et  $I_0^{(MA)}$  pour (c) et (d). (a) et (c) : les hypothèses d’application des coordonnées barycentriques généralisées sont vérifiées ; (b) et (d) : elles ne le sont pas.

Les figures 4.29-(a) et (c) montrent le critère de validation local  $D_n^{(f)}$  (équation (4.67)) après la première itération (ite\_1) et la dernière itération (ite\_fin) de l’algorithme de relaxation de l’étape de suivi pour un nœud  $S_n$  d’une image  $f$  d’une séquence d’images artificielle en fonction de la méthode de suivi que nous proposons : CMPT\_aN\_sPR (méthode 1), CMPT\_aN\_sMB (méthode 2), CMPT\_aN\_sCBG (méthode 3), CMPT\_aE\_sPR (méthode 4), CMPT\_aE\_sMB (méthode 5) et CMPT\_aE\_sCBG (méthode 6). Les résultats sont présentés pour les deux séquences d’images « tn1 » et « tgrzn2 » respectivement aux images 1 et 4 et aux nœuds 1 et 4. Dans ces deux cas, toutes les hypothèses d’application des coordonnées barycentriques généralisées sont vérifiées. La transformation  $T$  entre l’image de référence  $I_0$  et l’image  $I_{t-1}$  est affine. Celle entre  $I_{t-1}$  et  $I_t$  est nodale. Pour la séquence « tn1 »,  $T$  est l’identité (mais on obtiendrait les mêmes résultats avec une translation). Pour la séquence « tgrzn2 »,  $T$  est la composition d’une translation, d’une rotation et d’un zoom.

Les figures 4.29-(b) et (d) montrent le critère de validation locale moyen  $D_m$ , calculé pour tous les  $N_i$  nœuds internes du maillage (les tests ne sont pas effectués sur les nœuds du bord car les nœuds appartiennent aux sommets de leur polygone support) et toutes les  $F$

images de la séquence :

$$D_m = \frac{1}{N_i F} \sum_{f=1}^F \sum_{n=1}^{N_i} D_n^{(f)}, \quad (4.72)$$

après la première itération (ite\_1) et la dernière itération (ite\_fin) de l'algorithme de relaxation de l'étape de suivi en fonction de la méthode de suivi sPR, sMB ou sCBG. Ces résultats sont donnés pour quatre séquences d'images artificielles : « tgr3 », « tgz2 », « tgtrzn2 » et « tgtrzn2 ». Pour les trois premières séquences, les images sont déformées selon une transformation affine globale. Ceci n'est pas le cas pour la dernière séquence. Il s'agit d'une transformation affine par morceaux. Pour ces quatre séquences, les hypothèses d'application des coordonnées barycentriques généralisées ne sont donc pas exactement vérifiées.

## B) Analyse

- i) Lorsque les hypothèses d'application des coordonnées barycentriques généralisées sont vérifiées, la courbe en trait plein rouge correspondant à la séquence « tn1 » des deux figures 4.29-(a) et (c) montre que les trois méthodes (sPR, sMB et sCBG) de l'étape de suivi sont équivalentes. Pour l'apprentissage exact, l'algorithme de relaxation s'arrête en une itération (les courbes rouges en trait plein et en tirets sont confondues).

Par contre, les résultats sur la séquence « tgtrzn2 » (les images et le maillage étant déformés par une transformation affine globale) illustrent le fait qu'appliquer directement la sortie du prédictor est insuffisant. En effet, après la première itération, les écarts d'erreur entre la méthode de suivi sPR et les méthodes sMB et sCBG sont importants. Ceci est aussi vérifié et validé sur les données moyennes des figures 4.29-(b) et (d).

**On constate aussi que lorsque les hypothèses sont vérifiées, l'apprentissage exact donne de meilleurs résultats que l'apprentissage naïf.** Les courbes correspondant à l'apprentissage naïf, c'est-à-dire celles de la partie gauche des figures 4.29-(a) et (c), sont au-dessus de celles de la partie droite qui correspondent à l'apprentissage exact.

- 2i) Les résultats illustrés sur la figure 4.29 montrent que la qualité des méthodes sMB et sCBG est la même, ceci est vérifié dans le cas où le maillage est déformé selon une transformation affine globale. En effet, dans ce cas, les deux méthodes donnent analytiquement les mêmes résultats, ce qui est donc numériquement vérifié. Lorsqu'on n'est pas dans ce cas, ce qui se produit dans la séquence « tgtrzn2 » présentée sur les courbes bleues des figures 4.29-(b) et (d), en moyenne, l'écart est de l'ordre que  $10^{-3}$  pixel. Le critère  $D_m$  après la première itération de l'algorithme CMPT\_aN\_sMB est de 0.4270 pixel et celui pour l'algorithme CMPT\_aN\_sCBG est de 0.4255 pixel. **Au-tant dire qu'à l'ordre des déformations étudiées, les deux méthodes sont comparables.**
- 3i) Si les hypothèses d'application des coordonnées barycentriques généralisées ne sont pas exactement vérifiées (voir les courbes des figures 4.29-(b) et (d)), **on constate que l'apprentissage naïf donne de meilleurs résultats que l'apprentissage exact** : les courbes correspondant à l'apprentissage naïf, c'est-à-dire celles de la partie gauche de ces deux figures, sont en dessous de celles de la partie droite qui correspondent à l'apprentissage exact. Dans l'algorithme de relaxation, le nombre d'itérations nécessaire est donc supérieur pour l'apprentissage exact que naïf. On peut donc se demander ce qui est à préférer : garder un modèle mathématique exact et en moyenne avoir besoin

de plus d'itérations pour converger ou avoir un modèle d'approximation et converger plus rapidement... Ce point est de nouveau étudié dans la partie concernant les validations globales.

- 4i) Si l'on ne s'intéresse qu'aux courbes relatives à la valeur des critères à la fin des itérations de l'algorithme de relaxation, c'est-à-dire les courbes en tirets, on constate que toutes les méthodes sont équivalentes avec cependant des erreurs légèrement supérieures dans l'approche de suivi sPR, c'est-à-dire consistant à appliquer directement la sortie du prédicteur. **La relaxation est la source de la meilleure performance d'un algorithme itéré.**

#### 4.4.2.2 Validation globale sur des séquences d'images artificielles

##### A) Des chiffres, des courbes

La figure 4.30 montre les images 13 et 43 de la séquence d'images artificielle générée à partir de l'image de référence de  $I_0^{(F)}$ , le maillage déterminé par la vérité terrain (en bleu) et celui obtenu par divers algorithmes de suivi (NH avec un pas de recherche 0.1 pixel, NH avec un pas de recherche de 0.25 pixel, CMPT\_aN\_sCBG et CMPT\_aE\_sCBG).

Les courbes de la figure 4.31-(a) (respectivement 4.32-(a)) montrent, en fonction du numéro de l'image de la séquence artificielle « foreman\_artificiel » (respectivement « missa\_artificiel »), les valeurs du critère  $D_{moy}^{(f)}$  (équation (4.68)) pour les divers algorithmes. La courbe correspondant à l'algorithme NH avec un pas de recherche de 0.25 n'a pas été représentée. En effet, elle donne de moins bons résultats que tous les autres algorithmes et surchargeait le graphique.

Les courbes de la figure 4.31-(b) (respectivement 4.32-(b)) montrent, en fonction du numéro de l'image de la séquence artificielle « foreman\_artificiel » (respectivement « missa\_artificiel »), le nombre d'itérations de l'algorithme de relaxation de l'étape de suivi des méthodes CMPT\_aN\_sCBG et CMPT\_aE\_sCBG.

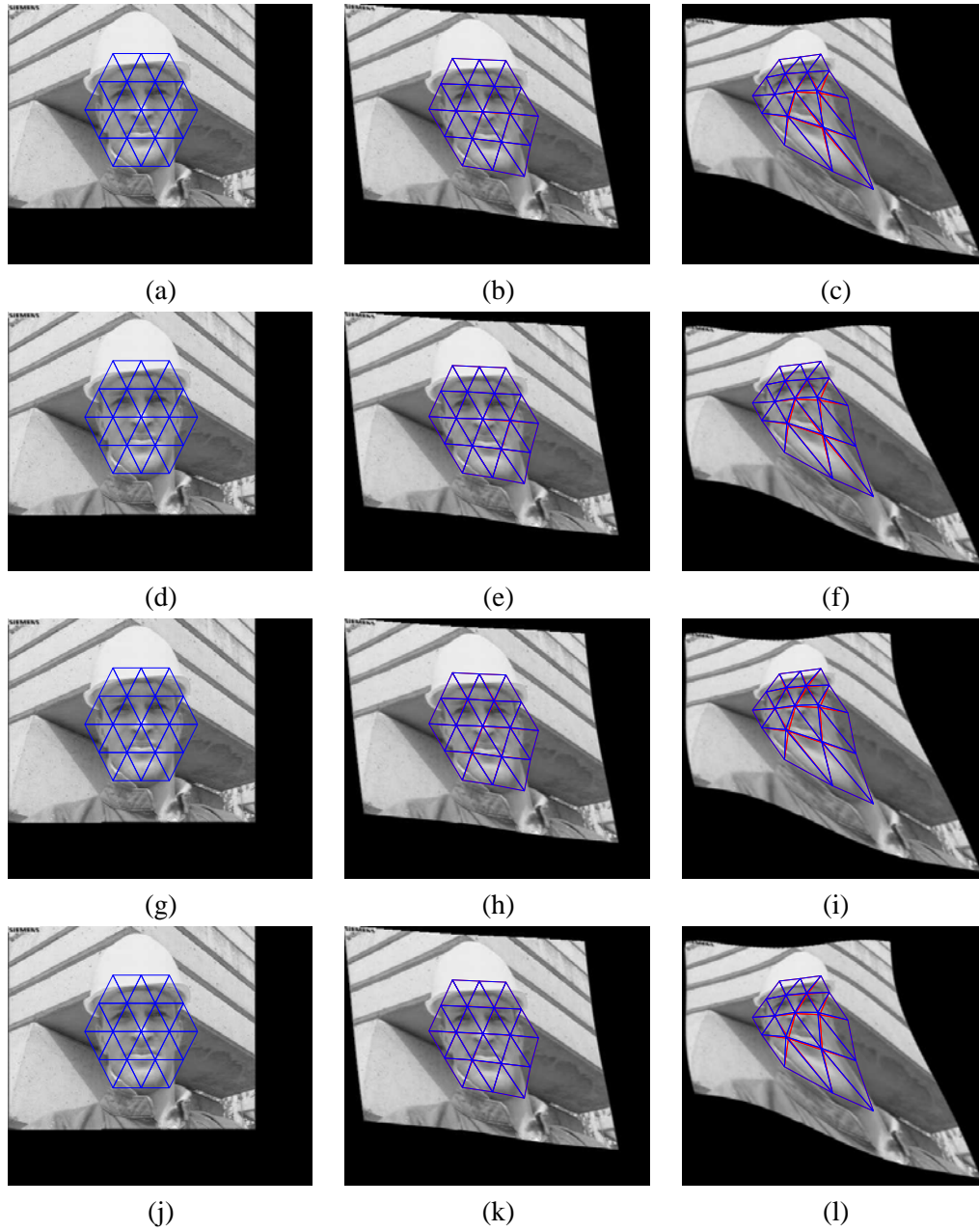


FIG. 4.30 – La séquence « foreman\_artificiel ». Colonne 1 : image de référence  $I_0^{(F)}$  et maillage d'apprentissage ; colonne 2 : image 13 ; colonne 3 : image 43. En rouge, la vérité terrain. En bleu, le maillage obtenu par divers algorithmes de suivi (remarquons que les deux maillages sont très proches) : ligne 1 : algorithme CMPT\_aN\_sCBG ; ligne 2 : algorithme CMPT\_aE\_sCBG ; ligne 3 : algorithme NH avec un pas de recherche de 0.25 ; ligne 4 : algorithme NH avec un pas de recherche de 0.1.



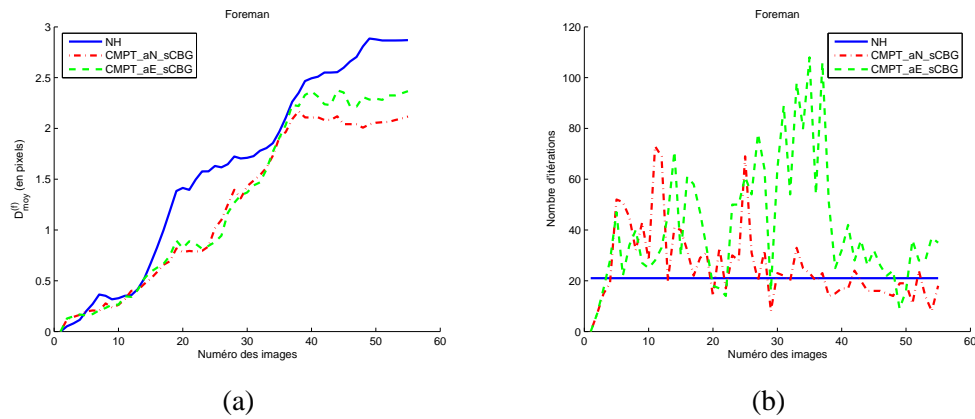


FIG. 4.31 – Suivi sur la séquence « foreman\_artificiel » présentée sur la figure 4.28 : (a) erreur moyenne à la vérité terrain pour les algorithmes NH avec un pas de 0.1 pixel, CMPT\_aN\_sCBG et CMPT\_aE\_sCBG ; (b) nombre d'itérations pour chacun de ces algorithmes.

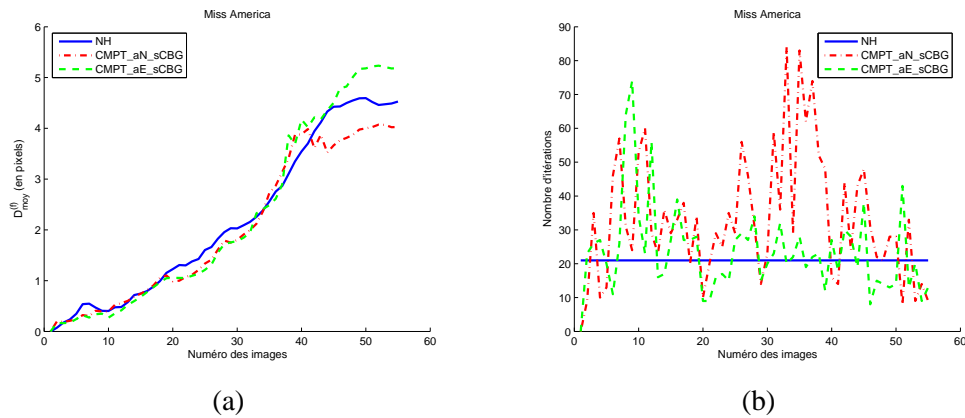


FIG. 4.32 – Suivi sur la séquence « missa\_artificiel » : (a) erreur moyenne à la vérité terrain pour les algorithmes NH avec un pas de 0.1 pixel, CMPT\_aN\_sCBG et CMPT\_aE\_sCBG ; (b) nombre d'itérations pour chacun de ces algorithmes.

## B) Analyse

- i) La figure 4.30 montre que, visuellement, toutes les méthodes de suivi donnent des résultats cohérents.
- 2i) Les figures 4.31-(a) et 4.32-(a) montrent que, même si le modèle de déformation des images n'est pas affine par morceaux, **notre méthode approche bien ces transformations non rigides**. En effet, pour la séquence « foreman\_artificiel », l'amplitude de déplacements des nœuds internes du maillage est comprise entre 0.0768 pixel et 5.3723 pixels et le critère de validation moyen  $D_{moy}^{(f)}$  maximal est de 2.37 pixels obtenu à l'image 44 pour l'algorithme CMPT\_aE\_sCBG, de 2.17 pixels obtenu à l'image 38 pour l'algorithme CMPT\_aN\_sCBG et de 2.88 pixels obtenu à l'image 49 pour l'algorithme NH.

Pour la séquence « missa\_artificiel », l'amplitude de déplacements des nœuds internes du maillage est comprise entre 0.1217 pixel et 6.7282 pixels et le critère de validation moyen  $D_{moy}^{(f)}$  maximal est de 5.23 pixels obtenu à l'image 52 pour l'algorithme CMPT\_aE\_sCBG, de 4.07 pixels obtenu à l'image 52 pour l'algorithme CMPT\_aN\_sCBG et de 4.59 pixels obtenu à l'image 50 pour l'algorithme NH.

- 3i) Le critère  $D_{moy}^{(f)}$  croît au fil des images d'une séquence aussi bien dans le cas de l'algorithme NH que nos méthodes CMPT. Ceci est certainement dû au cumul des erreurs de suivi tout au long de la séquence.
- 4i) Nos deux approches d'apprentissage naïf (aN) et exact (aE) donnent des résultats très similaires sur les débuts des séquences, après l'apprentissage naïf donne de meilleurs résultats que l'apprentissage exact. Les déformations des séquences peuvent en être à l'origine : « étirement » puis « contraction » de l'image de référence.
- 5i) Les figures 4.31-(b) et 4.32-(b) montrent qu'alors que pour la séquence « foreman\_artificiel », le nombre d'itérations moyen pour l'algorithme CMPT\_aN\_sCBG est de 25.76, celui pour l'algorithme CMPT\_aE\_sCBG est de 41.12. Cette observation n'est pas vérifiée pour la séquence « missa\_artificiel » avec un nombre d'itérations moyen de 33.03 pour l'algorithme CMPT\_aN\_sCBG et de 23.74 pour l'algorithme CMPT\_aE\_sCBG. Ceci est sûrement dû aux caractéristiques intrinsèques des textures différentes des deux images de référence.

La prédominance de l'apprentissage naïf ou exact reste encore à définir.

Il est à noter que la courbe sur le nombre d'itérations pour l'algorithme NH n'est pas vraiment comparable aux autres courbes. En effet, on ne parle pas exactement du même type d'itérations. Elle est fixée à 21 (= 3 itérations  $\times$  7 nœuds internes). On peut cependant remarquer qu'en moyenne, pour notre méthode donnant le moins d'itérations, CMPT\_aE\_sCBG pour « foreman\_artificiel » et CMPT\_aN\_sCBG pour « missa\_artificiel », ce nombre est comparable à celui de l'algorithme NH.

#### 4.4.2.3 Validation globale sur des séquences d'images réelles

##### A) Des chiffres, des courbes

La figure 4.33 (respectivement la figure 4.34) montre des résultats obtenus sur la séquence « Miss America » (respectivement « Foreman ») en utilisant l'algorithme NH avec un pas de recherche de 0.25 pixel (figure 4.33-(a)) (respectivement figures 4.34 (1-a), (1-b) et (1-c)) et nos algorithmes CMPT\_aN\_sCBG et CMPT\_aE\_sCBG (respectivement figures 4.33-(b) et (c)) (respectivement figures 4.34 (2-a), (2-b), (2-c) et (3-a), (3-b), (3-c)). Rappelons que pour une séquence d'images réelle, le mouvement des nœuds du bord du maillage n'est pas connu.

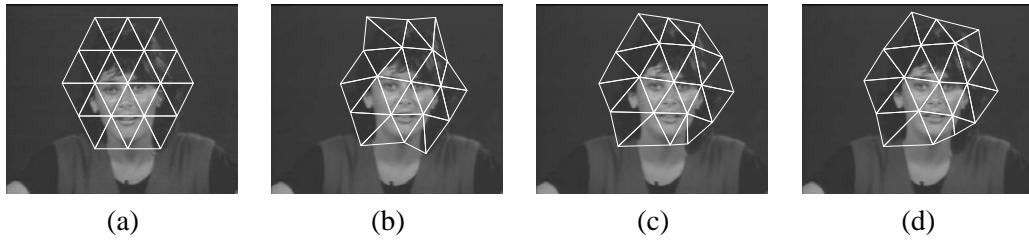


FIG. 4.33 – Séquence « Miss America ». (a) Image de référence et maillage initial. (b-c-d) Image numéro 78 et son maillage suivi par l'algorithme (a) NH avec un pas de recherche de 0.25 pixel. PRSB de 35.15 dB ; (b) CMPT\_aN\_sCBG. PRSB de 35.65 dB ; (c) CMPT\_aE\_sCBG. PRSB de 35.37 dB.

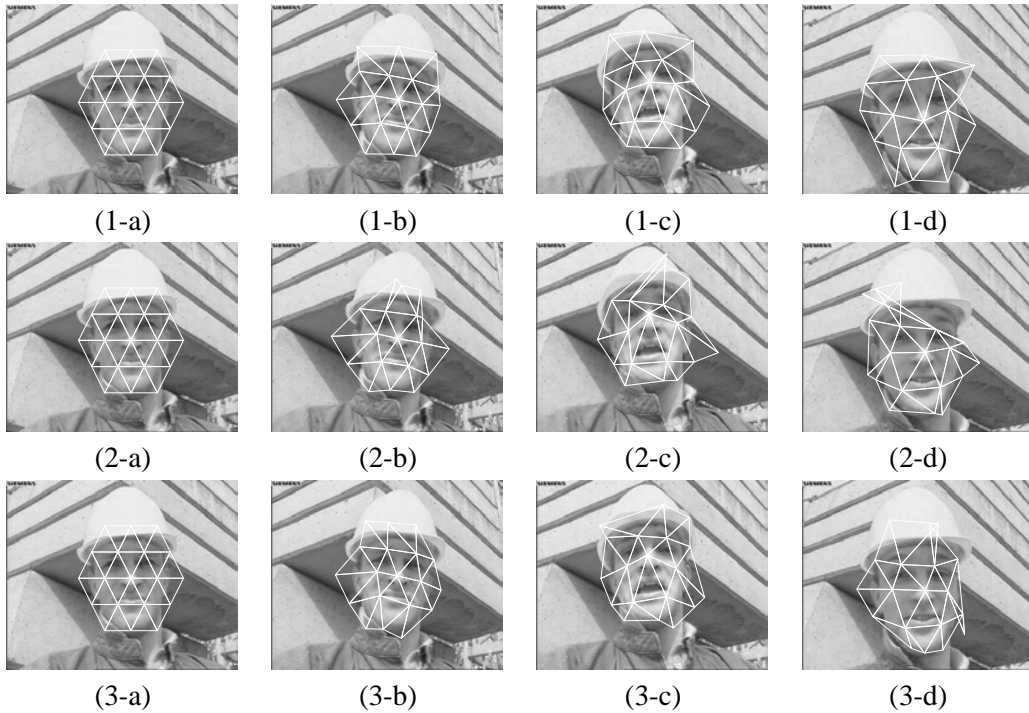


FIG. 4.34 – Séquence « Foreman ». Ligne 1 : maillage suivi par l'algorithme NH avec un pas de recherche de 0.25 pixel ; Ligne 2 : maillage suivi par l'algorithme CMPT\_aN\_sCBG ; Ligne 3 : maillage suivi par l'algorithme CMPT\_aE\_sCBG. (a) image de référence ; (b) image 40 ; (c) image 70 ; (d) image 100.

Les courbes de la figure 4.35 montrent, en fonction du numéro de l'image de la séquence, les valeurs du PRSB pour l'algorithme NH et nos deux méthodes.

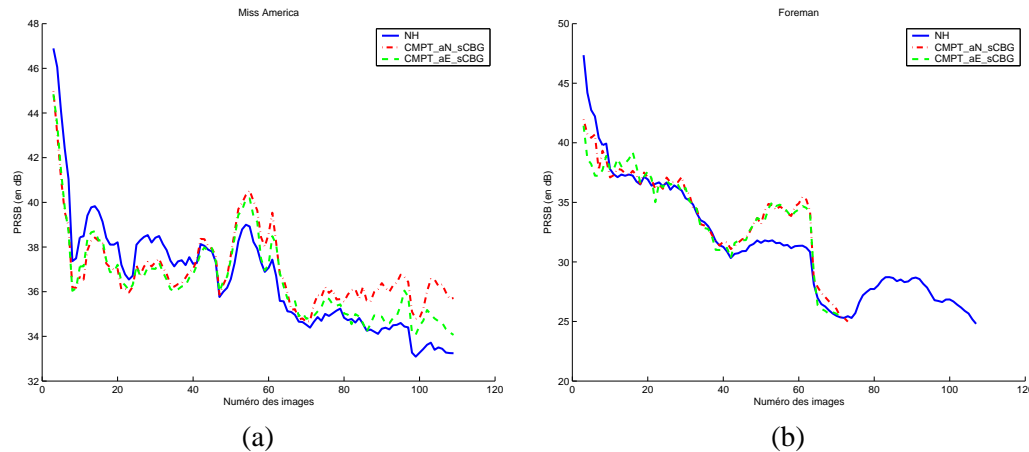


FIG. 4.35 – Comparaison entre les algorithmes NH, CMPT\_aN\_sCBG et CMPT\_aE\_sCBG sur la séquence d'images réelles (a) « Miss America » et (b) « Foreman ».

## B) Analyse

- i) Les courbes des figures 4.35-(a) et (b) montrent qu'en qualité de rendu visuel, les trois méthodes sont proches les unes des autres. La décroissance du PRSB est sûrement due au modèle de déformation choisi (la déformation d'un maillage) et au cumul des erreurs de suivi tout au long de la séquence.
- 2i) Les images de la figure 4.34 sur la séquence « Foreman » illustrent le fait que, contrairement à la séquence « Miss America » (voir figure 4.33), l'algorithme de suivi que nous proposons a ses limites. En effet, notre algorithme suit correctement le mouvement jusqu'aux images 73 et 71 pour respectivement les algorithmes CMPT\_aN\_sCBG et CMPT\_aE\_sCBG. Les figures 4.34-(2-c) et (3-c) montrent la position du maillage suivi sur l'image 70. Sur cette image, on peut voir que le chef de chantier a grandement ouvert la bouche et le maillage, surpris, commence à avoir du mal à suivre. Au-delà de ces images, un retournement de maille se produit (voir les figures 4.34-(2-d) et (3-d), l'illustration est sur l'image 100 de la séquence) et l'algorithme ne se stabilise pas ensuite. Ceci est essentiellement dû aux caractéristiques de la séquence d'images : mouvements très amples de la caméra et du chef de chantier, problème d'occultation non traité actuellement dans notre algorithme notamment lorsque le chef de chantier parle et tourne la tête, non détection des retournements de mailles, changements d'illuminations (ombre du casque). Enfin, les conditions d'application des coordonnées barycentriques généralisées sont largement non vérifiées.
- 3i) Les illustrations pour l'algorithme NH (voir les figures 4.34-(1-b), (1-c) et (1-d)) montrent un maillage globalement moins déformé. En effet, cet algorithme ne visite que trois fois chaque nœud dont les déplacements sont limités par la dimension de la fenêtre de recherche. Ceci n'est pas le cas dans notre algorithme de relaxation. On ne peut connaître à l'avance ni le nombre d'itérations ni le nœud traité pendant l'itération. Il peut donc arriver qu'un même nœud soit déplacé un grand nombre de fois conduisant ainsi à de grands déplacements. Une solution pour contrer ce problème serait de considérer une pénalité contraignant ainsi le maillage à rester conforme. Cette pénalité pourrait donc porter sur la compacité des mailles.

## 4.5 Bilan

Dans ce chapitre, nous avons présenté une nouvelle méthode de suivi (CMPT) d'objets déformables fondée sur un maillage triangulaire que nous avons validé aussi bien sur des séquences d'images artificielles que réelles et comparé à l'algorithme NH.

Le principal avantage de notre algorithme CMPT est d'être moins complexe, en temps de calcul, que l'algorithme NH. Il est regrettable qu'il ne donne que des résultats similaires à l'algorithme NH, qui est considéré comme un algorithme de référence. Il est donc intéressant de faire aussi bien que lui et cela plus efficacement.

Nous avons aussi détaillé avec précision toutes les briques de base de notre algorithme CMPT, à savoir une description géométrique et analytique :

- des transformations géométriques (globales et nodales) d'un maillage,
- de l'inverse d'une transformation nodale,
- de la composée d'une transformation globale et d'une transformation nodale, ainsi que son approximation en utilisant les coordonnées barycentriques généralisées.



## Chapitre 5

# Conclusion et perspectives

### 5.1 Bilan

Dans cette thèse, notre intérêt s'est porté sur le suivi d'objets déformables, au sein d'une séquence d'images, en utilisant l'outil maillage. Nous avons choisi non pas une représentation 3D de l'objet (via un maillage 3D par exemple), mais une représentation 2D de la projection de l'objet sur le plan image. Ainsi, nous avons étudié une représentation par :

- des caractéristiques visuelles, ici des points d'intérêt, reflet local de la texture de l'image,
- un maillage plan bidimensionnel qui n'est rien d'autre qu'un ensemble fini de points structurés et permet d'avoir une représentation compacte de l'objet tout au long du suivi. Ce choix a évidemment ajouté des contraintes de conservation de la topologie du maillage.

Les caractéristiques visuelles ne sont pas suivies indépendamment les unes des autres, mais l'ensemble des caractéristiques visuelles est suivi en formant un tout.

Nous avons travaillé sous l'hypothèse de l'invariance des niveaux de gris. Ainsi, l'étude de la modification de l'apparence nous a permis d'apprendre à prédire les déformations.

Dans ce mémoire de thèse, le premier axe de recherche sur lequel nous nous sommes penchés concernait l'amélioration de l'algorithme de suivi d'objets plans indéformables de Jurie et Dhome. Il a consisté à ajouter à cet algorithme une étape préliminaire de sélection de l'ensemble des points d'intérêt intervenant dans l'algorithme. Nous rappelons que cet algorithme est fondé sur un ensemble de points qui est la seule source d'information, donc la sélection judicieuse de ces points est indispensable en vue d'une amélioration. Cette sélection dépend nécessairement de la séquence d'images. Dans cette optique de sélection de points d'intérêt, nous avons utilisé de manière originale des méthodes de sélection d'un sous-ensemble provenant des statistiques et de l'algèbre linéaire.

Le second axe de recherche concernait une nouvelle méthode de suivi d'objets déformables en utilisant un maillage. Cette méthode reprend le principe de l'algorithme de Jurie et Dhome, à savoir une étape d'apprentissage hors ligne d'un prédicteur linéaire pour chaque nœud du maillage et une étape de suivi en ligne, fondé sur un algorithme de relaxation pour prendre en compte l'apport de chaque nœud. Cette généralisation de l'algorithme de suivi de Jurie et Dhome au cas de déformations d'un maillage, qui ne forment pas un groupe, a inévitablement impliqué la définition de toutes les briques de base de l'algorithme. En particulier, nous avons défini l'inverse d'une transformation nodale et, étant donné que l'image de la composition d'une transformation nodale et d'une transformation

globale d'un maillage n'est pas un maillage, nous avons proposé une nouvelle approximation de cette transformation et avons opté pour l'utilisation des coordonnées barycentriques généralisées en raison de leur propriété d'invariance par transformation affine globale.

Concernant, le premier axe de recherche, nous avons montré :

- qu'il n'existe pas de détecteur de points universel et que le meilleur ensemble de points d'intérêt dépend fortement de la séquence d'images traitée ;
- que la sélection d'un ensemble de points répartis uniformément est à privilégier. Cela permet d'être, par exemple, moins sensible au problème d'occultation ;
- qu'il ne faut pas prendre naïvement les meilleurs points d'intérêt au sens de leur critère de détection mais qu'il est aisé d'améliorer l'algorithme de suivi en sélectionnant un sous-ensemble de points, en donnant priorité à la méthode de sélection séquentielle par ajout (provenant des statistiques) (SFS\_C/SFS) par rapport à celle de *Golub et al.* (provenant de l'algèbre linéaire) (S\_C/S).

À l'issue de ce travail, nous recommandons donc de faire de la sélection en utilisant la méthode statistique et en répartissant uniformément les points.

Concernant, le second axe de recherche, nous avons montré que notre algorithme :

- est plus efficace que l'algorithme de Nakaya et Harashima qui effectue une recherche exhaustive des déplacements de tous les nœuds du maillage ;
- donne des résultats qui ont une précision équivalente à ceux de Nakaya et Harashima.

Comme nous l'avons vu, nous sommes frustrés par la qualité des résultats, ce qui nous conduit à poser deux questions :

- l'introduction d'une contrainte de conservation de la topologie du maillage serait-elle trop coûteuse ?
- existe-t-il une alternative au contraignant maillage 2D ?

La figure imposée « maillage » est certes « contraignante » mais finalement pas suffisamment « régularisante ».

## 5.2 Contributions

Rappelons succinctement nos contributions :

- 1) une variante de l'algorithme de suivi (JD) : §2.5.2 ;
- 2) une utilisation originale, dans le domaine du traitement d'images et de la vision, de méthodes statistique et algébrique de sélection d'un sous-ensemble : §3.4 ;
- 3) un nouvel algorithme de sélection d'un sous-ensemble de points d'intérêt fondé sur une sélection séquentielle par ajout des points d'intérêt : §3.4.1.4 ;
- 4) une caractérisation géométrique et analytique d'une transformation nodale et de son inverse : §4.1.5.3 et §4.1.5.4 ;
- 5) une nouvelle approximation de la composition d'une transformation globale et d'une transformation nodale d'un maillage mettant en œuvre les coordonnées barycentriques généralisées : §4.1.5.5 ;
- 6) un nouvel algorithme de suivi d'objets déformables utilisant un maillage : §4.3.



## 5.3 Perspectives

Au vu des résultats obtenus, nous sommes à même de proposer quelques perspectives, certaines regroupant les deux axes de recherche.

### Concernant la sélection d'un sous-ensemble de points d'intérêt

La sélection d'un sous-ensemble de points d'intérêt présentée pour l'algorithme de suivi d'objets indéformables de Jurie et Dhome doit pouvoir se généraliser pour notre algorithme de suivi. Quelques points particuliers devront être pris en compte, comme par exemple, le fait que l'on travaille sur des zones plus petites, donc moins riches en nombre de points d'intérêt. En outre, la transformée étudiée est une transformation nodale et on dénombre autant de sous-ensembles que de nœuds dans le maillage.

### Concernant le suivi par maillage

Dans la partie « résultats », nous avons vu que le suivi finit par échouer quand la séquence d'images évolue en s'éloignant trop de l'image de référence ou que les amplitudes des transformations sont excessives vis-à-vis de celles apprises. Des améliorations possibles consisteraient à :

- détecter les retournements de maille et les traiter a posteriori ou, au contraire, les prévenir en contraignant le maillage à ne pas se déformer trop fortement, par exemple en rajoutant un terme de régularisation,
- traiter les nœuds du bord selon leurs particularités. Typiquement, un nœud du bord est un sommet de son polygone support,
- mettre à jour, de temps en temps, les prédicteurs linéaires associés à chaque nœud tout au long du suivi pour qu'ils soient plus représentatifs du contexte actuel.

Une autre perspective serait d'aborder l'approximation de la composée d'une transformation nodale et d'une transformation globale d'un maillage par une approche analytique consistant par exemple à minimiser l'écart entre les arêtes discontinues de la composée exacte et un maillage déformé « normalement ».

Enfin, pour fournir un algorithme encore plus complet, il faudrait prendre en compte les problèmes d'occultation et les variations de niveaux de gris.

### Concernant l'analyse multi-résolution et les approches hiérarchiques

Pour améliorer la précision du suivi, nous pouvons imaginer faire du suivi hiérarchique, c'est-à-dire appliquer successivement le suivi pour des amplitudes décroissantes de déformations. Cela consiste à étendre la sélection des points d'intérêt afin de l'adapter au suivi hiérarchique. En effet, les points sélectionnés doivent être valides pour tous les niveaux de la hiérarchie.

Pour être résistant aux différentes amplitudes de déformation, une étude portant sur la multi-résolution de maillage permettrait d'approcher plus finement les déformations.

## 5.4 Les articles

Les travaux de cette thèse ont été publiés et présentés dans des conférences. Regroupons-les par thème :

- la sélection d'un sous-ensemble de points d'intérêt : [PCT07] et [PTC07] ;
- le suivi d'objets déformables par maillage : [PMC04], [PCM04] (premiers résultats), [PCM05b] et [PCM05a] (approche complète).

## Annexe A

# Rappel sur les polynômes orthogonaux discrets de Tchebychev : la théorie

Pour plus de détails, se reporter à l'ouvrage de Haralick et Shapiro [[HS93a], §8.8, p. 393 et Annexe B, p. 659] et à l'article de Haralick [Har84].

### A.1 Cas de la dimension 1

#### A.1.1 Une base de polynômes orthogonaux discrets de Tchebychev

Étant donné un ensemble de points  $J = \{-x_k, \dots, \{0\}, \dots, x_k\}$ , où :

$$\forall i = 1..k-1, x_{i+1} - x_i = cste = \begin{cases} x_1, & \text{si } \{0\} \in J \\ 2x_1, & \text{sinon} \end{cases},$$

la procédure itérative de construction de la base est la suivante :

$$\begin{cases} P_0(x) = 1, \\ P_i(x) = x^i + a_{i-1}x^{i-1} + \dots + a_1x + a_0, \end{cases} \quad (\text{A.1})$$

où  $P_i$  doit satisfaire à la relation d'orthogonalité :

$$\sum_{x_j \in J} P_k(x_j) P_i(x_j) = 0, \forall k = 0, \dots, i-1 \quad (\text{A.2})$$

On obtient :

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= x^2 - \frac{\mu_2}{\mu_0}, \\ P_3(x) &= x^3 - \frac{\mu_4}{\mu_2}x, \\ P_4(x) &= \frac{x^4 + (\mu_2\mu_4 - \mu_6)x^2 + (\mu_2\mu_6 - \mu_4^2)}{\mu_0\mu_4 - \mu_2}, \\ &\vdots \end{aligned} \quad (\text{A.3})$$

où :

$$\mu_k = \sum_{x_j \in J} x_j^k. \quad (\text{A.4})$$

### A.1.2 Le polynôme interpolateur de $f$

Le but est d'interpoler une fonction  $f$ , à l'aide d'un polynôme  $Q$  à une indéterminée, connaissant sa valeur aux points de  $J$ , soit  $\{f(-x_k), \dots, \{f(0)\}, \dots, f(x_k)\}$ .

À partir de la base de polynômes orthogonaux discrets de Tchebychev,  $Q$  s'écrit :

$$Q = \sum_{i=0}^{2k\{-1\}} a_i P_i \quad (\text{A.5})$$

( $2k$  si  $\{0\} \in J$ ,  $2k - 1$  sinon) et les  $a_i$  sont déterminés en écrivant :

$$Q(x_j) = \sum_{i=0}^{2k\{-1\}} a_i P_i(x_j) = f(x_j), \forall x_j \in J. \quad (\text{A.6})$$

La solution est unique et satisfait à :

$$a_i = \frac{\sum_{x_j \in J} P_i(x_j) f(x_j)}{\sum_{x_l \in J} (P_i(x_l))^2}. \quad (\text{A.7})$$

On en déduit que chaque coefficient  $a_i$  est obtenu par convolution discrète à savoir :

$$a_i = f|_J \otimes N_i, \quad (\text{A.8})$$

où  $N_i$  est un noyau de convolution défini par :

$$N_i(x_j) = \frac{P_i(x_j)}{\sum_{x_l \in J} (P_i(x_l))^2}. \quad (\text{A.9})$$

La dérivée de  $Q$  est donnée par :

$$Q' = \sum_{i=1}^{2k-\{1\}} a_i P'_i. \quad (\text{A.10})$$

## A.2 Cas de la dimension 2

Soient :

- $J_x = \{-x_n, \dots, \{0\}, \dots, x_n\}$  et  $J_y = \{-y_m, \dots, \{0\}, \dots, y_m\}$  un ensemble de points respectivement selon  $x$  et  $y$ ,
- $\{P_0, \dots, P_{2n\{-1\}}\}$  et  $\{Q_0, \dots, Q_{2m\{-1\}}\}$  une base de polynômes orthogonaux discrets à une indéterminée associée respectivement à  $J_x$  et  $J_y$ ,

alors les polynômes :

$$\{P_0(x)Q_0(y), P_1(x)Q_0(y), \dots, P_{2n\{-1\}}(x)Q_0(y), \dots, P_{2n\{-1\}}(x)Q_{2m\{-1\}}(y)\} \quad (\text{A.11})$$

constituent une base de polynômes orthogonaux discrets à deux indéterminées.

### A.3 Examples

$J_x = J_y$	“Bases”
$\{-1/2, 1/2\}$	$P_0(x) = 1, P_1(x) = x$
$\{-1, 0, 1\}$	$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2 - 2/3$
$\{-3/2, -1/2, 1/2, 3/2\}$	$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2 - 5/4,$ $P_3(x) = x^3 - 41/20 x$
$\{-2, -1, 0, 1, 2\}$	$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2 - 2,$ $P_3(x) = x^3 - 17/5, P_4(x) = x^4 + 3x^2 + 72/35$
$\{-1, 0, 1\} \times \{-1, 0, 1\}$	$\{1, x, y, x^2 - 2/3, xy, y^2 - 2/3, y(x^2 - 2/3),$ $x(y^2 - 2/3), (x^2 - 2/3)(y^2 - 2/3)\}$



## Annexe B

# Le mythe de l'opérateur « SOBEL »

Intrigués par les fondements mathématiques de l'opérateur de gradient de Sobel, nous avons tout simplement posé la question à son auteur. Irwin Sobel s'est fait le plaisir de nous répondre que :

- pareille question lui avait été posée près de 20 ans en arrière par Per-Erik Danielsson ;
- à cette occasion :
  - \* il avait « essayé » de reconstituer la genèse et l'historique de son opérateur ;
  - \* il avait peaufiné une note commencée plusieurs années auparavant ;
- le « tout » figure en « Introduction + Annexe » de l'article de P.E. Danielsson et O. Seger [[DS90], p. 348 et 376–379].

Avec les autorisations de toutes les parties concernées, nous joignons ces précieux documents à notre mémoire.

### B.1 La genèse d'un opérateur mythique

*The history of the “Sobel Operator” according to my best recollection is as follows :*

*In 1968, while a PhD candidate at the Stanford Artificial Intelligence Project I gave a talk, together with Gary Feldman (another graduate student and good friend of mine) on a relatively isotropic 3x3 gradient operator. This talk was presented at a time when the major piece of published work on computer vision was Larry Roberts' PhD Thesis from MIT wherein he defined a 2x2 gradient estimator then referred to as the “Roberts Cross” operator. I had previously thought up the operator via a line of reasoning presented in the accompanied short document and discussed it with Gary who enthusiastically proceeded to help me program it and test it. After doing so and satisfying ourselves that it gave at least visually desirable results, we presented these in a seminar at the Stanford Artificial Intelligence Project where we were employed as Research Assistants.*

*As this was an event that occurred more than 20 years ago my memory of it is somewhat foggy. Faculty that I most clearly remember in attendance at the seminar were Raj Reddy, John McCarthy and Arthur Samuels. I'm pretty sure that Peter Hart and/or Dick Duda and possibly Nils Nilsson from SRI were also there. Lester Earnest, project executive officer was most likely there. I'm pretty sure Karl Pingle who was employed as a project programmer and later wrote an edge follower incorporating the operator was there. Manfred Hueckel, a graduate student who later wrote a paper on a more robust and computationally expensive “edge detector”, was I think also there. Lynn Quam, Jay “Marty” Tenenbaum,*

*Gunnar Grape, Gil Falk, Dave Poole, and Phil Petit were other graduate students with the project and either were at the seminar or were working in such close proximity that they knew of the results.*

*My synopsis of what ensued was that Raj Reddy who was then teaching one of the first courses on Computer Vision coined the term “Sobel Operator” in contrast to the “Roberts Cross” and used it in his course. Subsequently Pingle published a paper (1969) describing it as part of his edge follower, and Duda and Hart mentioned it in their book.*

## B.2 Les fondements mathématiques de l'opérateur « SOBEL »

### An Isotropic 3x3 Image Gradient Operator

by Irwin Sobel

*We would like to document the derivation of a simple, computationally efficient, gradient operator which we developed in 1968. This operator has been frequently used and referenced since that time. The earliest description of this operator in the computer vision literature is [1], although it has been more widely popularized by its appearance in [2]. Horn [3] defines this operator and references 4 numerical analysis texts [4-7] with the statement :*

*“Numerical analysis [4-7] teaches us that for certain classes of surfaces an even better estimate is obtained using a weighted average of three such central differences... These expressions produce excellent estimates for the components of the gradient of the central point.”*

*The motivation to develop it was to get an efficiently computable gradient estimate which would be more isotropic than the then popular “Roberts Cross” operator [8]. The principle involved is that of estimating the gradient of a digitized picture at a point by the vector summation of the 4 possible simple central gradient estimates obtainable in a 3x3 neighborhood. The vector summation operation provides an averaging over directions-of-measurement of the gradient. If the density function was truly planar over the neighborhood all 4 gradients would have the same value. Any differences are deviations from local planarity of the function over the neighborhood. The intent here was to extract the direction of the “best” plane although no attempt was made to make this rigorous.*

*To be more specific, we will refer here to the image function as a “density” function. (It could just as well be an “intensity” function - the difference depends on the physical nature of the image source.) For a 3x3 neighborhood each simple central gradient estimate is a vector sum of a pair of orthogonal vectors. Each orthogonal vector is a directional derivative estimate multiplied by a unit vector specifying the derivative's direction. The vector sum of these 4 simple gradient estimates amounts to a vector sum of the 8 directional derivative vectors.*

*Thus for a point on a Cartesian grid and its eight neighbors having density values as shown*

<i>a</i>	<i>b</i>	<i>c</i>
<i>d</i>	<i>e</i>	<i>f</i>
<i>g</i>	<i>h</i>	<i>i</i>

*we define the magnitude of the directional derivative estimate vector 'g' for a given neighbor as*

$$|g| = \frac{\langle \text{density difference} \rangle}{\langle \text{distance to neighbor} \rangle}$$



The direction of 'g' will be given by the unit vector to the appropriate neighbor. Notice that the neighbors group into antipodal pairs : (a,i) (b,h) (c,g) (f,d). Vector summing derivative estimates within each pair causes all the "e" values to cancel leaving the following vector sum for our gradient estimate :

$$G = \frac{c-g}{4} * \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{a-i}{4} * \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{b-h}{2} * \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \frac{f-d}{2} * \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

the resultant vector being :

$$G = \begin{bmatrix} \frac{c-g-a+i}{4} + \frac{f-d}{2} \\ \frac{c-g+a-i}{4} + \frac{b-h}{2} \end{bmatrix}$$

Notice that the square root fortuitously drops out of the formula. If this were to be metrically correct we should divide result by 4 to get the average gradient. However, since these operations are typically done in fixed point on small integers and division loses low order significant bits, it is convenient rather to scale the vector by 4, thereby replacing the "divide by 4" (doubleshift right) with a "multiply by 4" (doubleshift left) which will preserve low order bits. This leaves us with an estimate which is 16 times as large as the average gradient. The resultant formula is :

$$G' = 4G = \begin{bmatrix} c - g - a + i + 2(f - d) \\ c - g + a - i + 2(b - h) \end{bmatrix}$$

It is useful to express this as weighted density summations using the following weighting functions for x and y components :

-1	0	1	1	2	1
-2	0	2	0	0	0
-1	0	1	-1	-2	-1
x-component			y-component		

This algorithm was used as an edgepoint detector in the 1968 vision system [2] at the Stanford Artificial Intelligence Laboratory wherein a point was considered an edgepoint if and only if

$$|G'|^2 > T$$

where  $T$  was a previously chosen threshold. For this purpose it proved an economical alternative to the more robust, but computationally expensive "Hueckel operator" [9].

### References :

- [1] Pingle, K.K., "Visual Perception by a Computer", in *Automatic Interpretation and Classification of Images*, A. Grasselli (Ed.), Academic Press, New York, 1969, pp. 277-284.
- [2] Duda, R.O. and Hart, P.E., pp. 271-272 in *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973
- [3] Horn, B.K.P., "Hill Shading and the Reflectance Map", Originally in *Proc. DARPA Workshop in Image Understanding*, Apr 24-25, 1979, p. 85 Science Applications Inc. Report SAI-80-895-WA ; Later in *Geo- Processing 2*(1982) p. 74, Elsevier Scientific Publishing, Amsterdam.
- [4] Conte, D. and de Boor, C., *Elementary Numerical Analysis*, 1972, New York : McGraw Hill.
- [5] Hamming, R.W., *Numerical Methods for Scientists and Engineers*, 1962, New York : McGraw Hill.

[6] Richtmeyer, R.D. and Morton, K.W., *Difference Methods for Initial-Value Problems*, New York : John Wiley pp.136-143.

[7] Hildebrand, F.B., *Introduction to Numerical Analysis*, 1956, 1974 New York : McGraw Hill.

[8] Roberts, L. G., "Machine Perception of Three-Dimensional Solids," in *Optical and Electro-Optical Information Processing*, pp. 159-197, J. T. Tippett, et al., (Ed.'s), MIT Press, Cambridge, Mass., 1965.

[9] Hueckel, M.H., "An Operator which Locates Edges in Digitized Pictures" in *Journal of the Association of Computing Machinery*, Vol.18, No. 1, January 1971, pp. 113-125.

## Annexe C

# Courbure d'une surface

### C.1 Les divers composants de base

Soient :

- 1)  $\mathcal{E}_3$  l'espace affine euclidien, normé, canonique, associé à l'espace vectoriel  $(\mathbb{R}^3, ||.||_2)$ .  $\mathcal{E}_3$  est supposé muni d'un repère orthonormé direct  $(O; (\mathbf{i}, \mathbf{j}, \mathbf{k}))$  ;
- 2)  $\Sigma$ , une surface paramétrée cartésienne de classe  $\mathcal{C}^k$ ,  $k \geq 2$ , définie par le couple  $(\Omega, F)$ , où :
  - $\Omega$  est un ouvert connexe de  $\mathbb{R}^2$
  - $F$  est une application de classe  $\mathcal{C}^{k(\geq 2)}$  sur  $\Omega$  :

$$\left\{ \begin{array}{ll} F : \Omega \subset \mathbb{R}^2 & \rightarrow (\mathcal{E}_3, \mathbb{R}^3) \\ [x \ y]^T & \mapsto F(x, y) = O + x\mathbf{i} + y\mathbf{j} + f(x, y)\mathbf{k} ; \\ & \stackrel{n}{=} [x \ y \ f(x, y)]^T \end{array} \right. ; \quad (\text{C.1})$$

- 3)  $M$ , un point générique de  $\Sigma$  paramétré par  $(x, y)$  et  $m_M = F(x, y)$  son correspondant dans  $\mathcal{E}_3$  ;
- 4)  $T_M$ , le plan tangent en  $M$  à  $\Sigma$ , un sous-espace vectoriel de dimension 2 de  $\mathbb{R}^3$ .  $T_M$  est le plan de  $\mathcal{E}_3$  passant par  $m_M$  associé ;
- 5)  $\mathcal{N}_M$ , le vecteur normal (unitaire) en  $M$  à  $\Sigma$  donc à  $T_M$ .

Alors :

- 1) la matrice jacobienne  $J_F$  associée à  $F$  est :

$$\forall (x, y) \in \Omega, J_F(x, y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x}(x, y) \stackrel{n}{=} p(x, y) & \frac{\partial f}{\partial y}(x, y) \stackrel{n}{=} q(x, y) \end{bmatrix}.$$

$\text{rang}(J_F(x, y)) = 2$ , ce qui signifie que tout point  $M(x, y)$  est régulier ;

- 2)  $\forall (x, y) \in \Omega$ ,

$$\frac{\partial F}{\partial x}(x, y) \stackrel{n}{=} F'_x(x, y) \equiv \begin{bmatrix} 1 \\ 0 \\ p(x, y) \end{bmatrix} \quad \text{et} \quad \frac{\partial F}{\partial y}(x, y) \stackrel{n}{=} F'_y(x, y) \equiv \begin{bmatrix} 0 \\ 1 \\ q(x, y) \end{bmatrix} ;$$

3)  $\forall (x, y) \in \Omega$ ,

$$\frac{\partial^2 F}{\partial x^2}(x, y) \stackrel{n}{=} F''_{x^2}(x, y) \equiv \begin{bmatrix} 0 \\ 0 \\ \frac{\partial^2 f}{\partial x^2}(x, y) \stackrel{n}{=} r(x, y) \end{bmatrix},$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = \frac{\partial^2 F}{\partial y \partial x}(x, y) \stackrel{n}{=} F''_{xy}(x, y) \equiv \begin{bmatrix} 0 \\ 0 \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) \stackrel{n}{=} s(x, y) \end{bmatrix},$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) \stackrel{n}{=} F''_{y^2}(x, y) \equiv \begin{bmatrix} 0 \\ 0 \\ \frac{\partial^2 f}{\partial y^2}(x, y) \stackrel{n}{=} t(x, y) \end{bmatrix};$$

4)  $T_M = Vect(\{F'_x(x, y), F'_y(x, y)\})$ ;

5)

$$\begin{cases} \mathbf{N} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3 \\ [x \ y]^T \mapsto \mathbf{N}(x, y) = \frac{F'_x(x, y) \wedge F'_y(x, y)}{\tilde{H}(x, y)} \end{cases}, \quad (\text{C.2})$$

où :

$$\begin{cases} \tilde{H} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R} \\ [x \ y]^T \mapsto \tilde{H}(x, y) = \|F'_x(x, y) \wedge F'_y(x, y)\|_2 \end{cases}. \quad (\text{C.3})$$

$\mathbf{N}$  et  $\tilde{H}$  sont de classe  $\mathcal{C}^{k-1}$ ,  $k \geq 2$ ;

6) En posant :

$$\begin{cases} \tilde{E} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R} \\ [x \ y]^T \mapsto \tilde{E}(x, y) = \|F'_x(x, y)\|_2^2 \end{cases},$$

$$\begin{cases} \tilde{F} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R} \\ [x \ y]^T \mapsto \tilde{F}(x, y) = (F'_x(x, y) | F'_y(x, y)) \end{cases},$$

$$\begin{cases} \tilde{G} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R} \\ [x \ y]^T \mapsto \tilde{G}(x, y) = \|F'_y(x, y)\|_2^2 \end{cases},$$

d'après un résultat attribué à Lagrange :

$$\forall (x, y) \in \Omega, \tilde{H}(x, y) = \sqrt{\tilde{E}(x, y)\tilde{G}(x, y) - (\tilde{F}(x, y))^2}$$

et par suite,  $\forall (x, y) \in \Omega$  :

$$\tilde{E}(x, y) = 1 + (p(x, y))^2, \quad \tilde{F}(x, y) = p(x, y)q(x, y), \quad \tilde{G}(x, y) = 1 + (q(x, y))^2,$$

$$\tilde{H}(x, y) = \sqrt{1 + (p(x, y))^2 + (q(x, y))^2},$$

$$\mathbf{N}(x, y) = \frac{1}{\tilde{H}(x, y)} \begin{bmatrix} -p(x, y) & -q(x, y) & 1 \end{bmatrix}^T.$$

## C.2 L'endomorphisme de Weingarten

### C.2.1 Définition

$$\forall (x, y) \in \Omega, \begin{cases} F'(x, y) (\equiv J_F(x, y)) \in \mathbb{R}^{3 \times 2} (\equiv \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)) \\ \mathbf{N}'(x, y) (\equiv J_{\mathbf{N}}(x, y)) \in \mathbb{R}^{3 \times 2} (\equiv \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)) \end{cases} \quad \text{prennent leurs va-}$$

leurs dans  $T_M$ , sous-espace vectoriel de  $\mathbb{R}^3$ , de dimension 2. En effet :

$$\forall \phi = [\phi_x \ \phi_y]^T \in \mathbb{R}^2, \begin{cases} F'(x, y)\phi = J_F(x, y)\phi = \phi_x F'_x(x, y) + \phi_y F'_y(x, y) \\ \mathbf{N}'(x, y)\phi = J_{\mathbf{N}}(x, y)\phi = \phi_x \mathbf{N}'_x(x, y) + \phi_y \mathbf{N}'_y(x, y) \end{cases},$$

où  $(\mathbf{N}'_x(x, y), \mathbf{N}'_y(x, y)) \in T_M$  (voir [[RDO98b], p. 185]).

En conséquence :

$$\begin{cases} \text{Im}(J_F(x, y)) \subset T_M \\ \text{rang}(J_F(x, y)) = 2 \end{cases} \Rightarrow \text{Im}(J_F(x, y)) = T_M.$$

Cette propriété permet de définir :

$$\begin{cases} G_M : T_M \subset \mathbb{R}^3 & \rightarrow \mathbb{R}^2 \\ [u \ v \ w]^T & \mapsto \phi = [\phi_x \ \phi_y]^T = G_M([u \ v \ w]^T) \end{cases},$$

tel que  $F'(x, y)\phi = J_F(x, y)\phi = [u \ v \ w]^T$ .

L'endomorphisme de Weingarten, noté  $w_M$ , associé au point  $M(x, y)$ , est alors défini par :

$$\begin{cases} w_M : T_M \subset \mathbb{R}^3 & \rightarrow T_M \subset \mathbb{R}^3 \\ [u \ v \ w]^T & \mapsto w_M([u \ v \ w]^T) = -(\mathbf{N}'(x, y) \circ G_M)[u \ v \ w]^T \end{cases} \quad (\text{C.4})$$

---

#### Remarque

$w_M$  s'identifie à la matrice  $-J_{\mathbf{N}}(x, y)(J_F(x, y))^+$ .

---

### C.2.2 Propriétés

- i)  $w_M$  est symétrique [[RDO98b], proposition II, p. 187],
- 2i)  $w_M$  est donc diagonalisable dans une base orthonormale formée des vecteurs propres unitaires associés aux deux valeurs propres réelles  $\lambda_1(M)$  et  $\lambda_2(M)$  distinctes ou confondues.

## C.3 Directions principales

Pour tout point  $M(x, y) \in \Sigma$ , les directions principales sont les vecteurs propres  $u_1(M) \in T_M$ ,  $u_2(M) \in T_M$  associés respectivement à  $\lambda_1(M)$ ,  $\lambda_2(M)$  dans le cas où  $\lambda_1(M) \neq \lambda_2(M)$ .

---

#### Remarque

Dans le cas où  $\lambda_1(M) = \lambda_2(M)$ , tout vecteur de  $T_M$  définit une direction principale.  $M$  est alors appelé "ombilic" de  $\Sigma$ .

---

## C.4 Courbures moyenne et gaussienne

### C.4.1 Les composants introduits par Gauss

Définissons (voir [[RDO98b], proposition II, p. 187]) :

$$\left\{ \begin{array}{ll} \tilde{L} : \Omega \subset \mathbb{R}^2 & \rightarrow \mathbb{R} \\ [x \ y]^T & \mapsto \tilde{L}(x, y) = (F''_{xx}(x, y)|\mathbf{N}(x, y)) = - (F'_x(x, y)|\mathbf{N}'_x(x, y)) \end{array} \right. ,$$

$$\left\{ \begin{array}{ll} \tilde{M} : \Omega \subset \mathbb{R}^2 & \rightarrow \mathbb{R} \\ [x \ y]^T & \mapsto \tilde{M}(x, y) = (F''_{xy}(x, y)|\mathbf{N}(x, y)) = - (F'_x(x, y)|\mathbf{N}'_y(x, y)) \\ & = - (F'_y(x, y)|\mathbf{N}'_x(x, y)) \end{array} \right. ,$$

$$\left\{ \begin{array}{ll} \tilde{N} : \Omega \subset \mathbb{R}^2 & \rightarrow \mathbb{R} \\ [x \ y]^T & \mapsto \tilde{N}(x, y) = (F''_{yy}(x, y)|\mathbf{N}(x, y)) = - (F'_y(x, y)|\mathbf{N}'_y(x, y)) \end{array} \right. .$$

Alors :

$$\forall (x, y) \in \Omega, \quad \tilde{L}(x, y) = \frac{r(x, y)}{\tilde{H}(x, y)}, \quad \tilde{M}(x, y) = \frac{s(x, y)}{\tilde{H}(x, y)} \quad \text{et} \quad \tilde{N}(x, y) = \frac{t(x, y)}{\tilde{H}(x, y)}.$$

On montre que les valeurs propres  $\lambda_1(M)$  et  $\lambda_2(M)$  de  $w_M$  sont les racines de l'équation (4) de Ramis [[RDO98b], éq. (4), p. 198] (nous oublions l'argument  $(x, y)$ ) :

$$(\tilde{E}\tilde{G} - \tilde{F}^2) \lambda^2 - (\tilde{G}\tilde{L} + \tilde{E}\tilde{N} - 2\tilde{F}\tilde{M}) \lambda + (\tilde{L}\tilde{N} - \tilde{M}^2) = 0 \quad (\text{C.5})$$

### C.4.2 Courbure moyenne : $\kappa_m(M)$

C'est tout simplement la moyenne arithmétique de  $\lambda_1(M)$  et  $\lambda_2(M)$ , à savoir :

$$\begin{aligned} \kappa_m(M) &= \frac{1}{2} \text{Tr}(w_M) \\ &= \frac{\lambda_1(M) + \lambda_2(M)}{2} \\ &= \frac{\tilde{G}\tilde{L} + \tilde{E}\tilde{N} - 2\tilde{F}\tilde{M}}{2\tilde{H}^2} \\ &= \frac{(1 + q^2)r + (1 + p^2)t - 2pqs}{2(1 + p^2 + q^2)^{3/2}} \end{aligned} \quad (\text{C.6})$$

### C.4.3 Courbure gaussienne : $\kappa_g(M)$

C'est tout simplement le carré de la moyenne géométrique de  $\lambda_1(M)$  et  $\lambda_2(M)$ , à savoir :

$$\begin{aligned} \kappa_g(M) &= \text{Det}(w_M) \\ &= \lambda_1(M)\lambda_2(M) \\ &= \frac{\tilde{L}\tilde{N} - 2\tilde{M}^2}{\tilde{H}^2} \\ &= \frac{rt - s^2}{(1 + p^2 + q^2)^2} \end{aligned} \quad (\text{C.7})$$

**Remarques**

- 1) Selon que  $\kappa_g(M)$  est strictement négative, respectivement nulle, respectivement strictement positive,  $M$  est dit hyperbolique, respectivement parabolique, respectivement elliptique.
- 2) Le numérateur de la courbure gaussienne  $\kappa_g(M)$  n'est autre que le Hessien (déterminant de la matrice Hessienne  $\nabla^2 f(x, y)$ ) :

$$\nabla^2 f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{bmatrix}$$

- 3) Le dénominateur de  $\kappa_g(M)$  vaut, quant à lui,  $\| [1 \quad (\nabla f(x, y))^T]^T \|^2$ . Comme dans le cas de la courbure d'une courbe plane, la courbure gaussienne fait intervenir la dérivée première de  $f$ . Il est donc **impossible** d'en faire abstraction comme se le permettent certains auteurs (cf. Harris [[HS88], p. 149]).
- 4) Le signe de  $\kappa_g(M)$  ne dépend que du numérateur donc de  $\text{Det}(\nabla^2 f(x, y))$  et, en conséquence,  $M$  est un point :
  - parabolique si  $\begin{cases} \lambda_{\min}(\nabla^2 f(x, y)) = 0 \\ \text{ou} \\ \lambda_{\max}(\nabla^2 f(x, y)) = 0 \end{cases}$ ,
  - elliptique si  $\begin{cases} \lambda_{\min}(\nabla^2 f(x, y)) \neq 0 \\ \lambda_{\max}(\nabla^2 f(x, y)) \neq 0 \\ \text{signe}(\lambda_{\min}(\nabla^2 f(x, y))) = \text{signe}(\lambda_{\max}(\nabla^2 f(x, y))) \end{cases}$ ,
  - hyperbolique si  $\begin{cases} \text{signe}(\lambda_{\min}(\nabla^2 f(x, y))) = -1 \\ \text{signe}(\lambda_{\max}(\nabla^2 f(x, y))) = +1 \end{cases}$ .
- 5) Il est **possible** de comparer les courbures gaussiennes de tous les points  $M(x, y)$  satisfaisant à  $\|\nabla f(x, y)\| = \text{constante}$ , en comparant les déterminants de leur matrice Hessienne  $\nabla^2 f(x, y)$  : c'est **le seul cadre qui donne sens** aux considérations faites autour de  $\nabla^2 f(x, y)$ .

**C.5 Courbures normale et principales****C.5.1 Courbure normale**

La courbure normale associée à une direction  $t_M$  tangente en  $M(x, y)$  à  $\Sigma$  a pour valeur absolue la courbure ( $\kappa_c$ ) en  $M$  de la courbe plane ( $\Gamma$ ) résultant de l'intersection de  $\Sigma$  avec le plan défini par son vecteur normal ( $t_M \wedge \mathbf{N}_M$ ) (voir [[RDO98b], proposition I, p. 195]).

**C.5.2 Courbures principales**

Les courbures principales de  $\Sigma$  en  $M(x, y)$  sont les courbures normales associées aux directions principales de  $\Sigma$  en  $M$  : ce sont  $\lambda_1(M)$  et  $\lambda_2(M)$  et elles correspondent aux extrema des courbures normales de  $\Sigma$  en  $M$  (voir [[RDO98b], théorème II, p. 195]).

---

**Remarques**

- 1) Il n'existe pas de lien direct entre directions principales et vecteurs propres de la matrice Hessienne  $\nabla^2 f(x, y)$  !
  - 2) Il n'existe pas de lien direct entre courbures principales et valeurs propres de  $\nabla^2 f(x, y)$  !
-



# Bibliographie

- [All00] P. Alliez. *Étude de la représentation géométrique et textuelle de scènes 3-D pour les services de visualisation dans un contexte télécommunicant*. Thèse de Doctorat, École Nationale Supérieure des Télécommunications, 2000.
- [AR75] N. Ahmed and K.R. Rao. *Orthogonal transforms for digital signal processing*. Springer-Verlag, 1975.
- [AT97a] Y. Altunbasak and A.M. Tekalp. Occlusion-adaptive content-based 2-D mesh design and forward tracking. *IEEE Transactions on Image Processing*, 6(9) :1270–1280, 1997.
- [AT97b] Yucel Altunbasak and A. Murat Tekalp. Closed-form connectivity-preserving solutions for motion compensation using 2-D meshes. *IEEE Transactions on Image Processing*, 6(9) :1255–1269, 1997.
- [Bal05] R. Balter. *Construction d'un maillage 3-D évolutif et scalable pour le codage vidéo*. Thèse de doctorat de l'Université de Rennes I, 2005.
- [Bau00] A. Baumberg. Reliable feature matching across widely separated views. *Computer Vision and Pattern Recognition*, 1 :774–781, 2000.
- [Bea76] P. Beaudet. Context dependent interpolation. *Image Science Mathematics Symposium*, pages 234–236, 1976.
- [Bea78] P. Beaudet. Rotationally invariant image operators. *International Conference on Pattern Recognition*, pages 579–583, 1978.
- [Bec73] P. Beckmann. *Orthogonal Polynomials for Engineering and Physicists*. The Golem Press, 1973.
- [Bjö96] A. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia (PA), 1996.
- [BL98] L. Bretzner and T. Lindeberg. Feature tracking with automatic selection of spatial scales. Technical report, KTH Kungliga Tekniska högskolan - Royal Institute of Technology, 1998.
- [BL02] M. Brown and D.G. Lowe. Invariant features from interest point groups. In *British Machine Vision Conference*, pages 656–665, 2002.
- [BM01] S. Baker and I. Matthews. Equivalence and efficiency of image alignment algorithms. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1090–1097, 2001.
- [BM04] S. Baker and I. Matthews. Lucas-Kanade 20 years on : A unifying framework. *International Journal of Computer Vision*, 56(3) :221–255, 2004.
- [Boo89] F.L. Bookstein. Principal warps : Thin-plate splines and the decomposition of deformations. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 11(6), pages 567–585, 1989.

- [Bra86] R. Bracewell. *The Fourier transform and its applications*. McGraw-Hill International Editions, second revised edition, 1986.
- [Buh00] M.D. Buhmann. Radial basis function. *Acta Numerica*, pages 1–38, 2000.
- [BZ04] A. Bartoli and A. Zisserman. Direct estimation of non rigid registration. In *British Machine Vision Conference*, volume 2, pages 899–908, 2004.
- [CET01] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 23(6) :681–685, 2001.
- [Cha94] V. Charvillat. *Un algorithme d’optimisation combinatoire pour l’appariement d’un couple d’images stéréoscopiques : intégration des algorithmes génétiques, du recuit simulé et de la recherche tabou*. Rapport de DEA, ENSEEIHT-Toulouse, 1994.
- [Che87] L.P. Chew. Constrained Delaunay triangulations. In *Proceedings of the Annual Conference on Computational Geometry*, ACM, pages 215–222, 1987.
- [Cia85] P.G. Ciarlet. *Introduction à l’analyse numérique matricielle et à l’optimisation*. Masson, 1985.
- [CLL05] R.T. Collins, Y. Liu, and M. Leordeanu. Online selection of discriminative tracking features. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 27(10) :1631–1643, 2005.
- [dBvKOS00] M. de Berg, M. van Kerveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry : Algorithms and Applications*. Springer-Verlag, 2000.
- [DC99] F. Dellaert and R. Collins. Fast image-based tracking by selective pixel integration. In *Proceedings of International Conference on Computer Vision Workshop on Frame-Rate Applications*, pages 1–22, 1999.
- [Dem02] L. Demaret. *Étude de la scalabilité et de la représentation d’images fixes par maillages hiérarchiques exploitant les éléments finis et les ondelettes bi-dimensionnelles : Application au codage vidéo*. Thèse de doctorat de l’Université de Rennes I, 2002.
- [Dou04] M. Douze. *Estimation d’homographies inter-images : cas des mosaïques et du suivi en temps réel, applications en réalité augmentée*. Thèse de doctorat, INPT, 2004.
- [DS83] J. E. Dennis and R.B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Prentice Hall, Englewood Cliffs, NJ, 1983.
- [DS90] P. Danielsson and O. Seger. Generalized and separable Sobel operators. In H. Freeman, editor, *Machine Vision for Three-Dimensional Scenes*, pages 347–379. Academic Press, 1990.
- [Duc76] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. In *RAIRO Analyse Numérique*, volume 10, pages 5–12, 1976.
- [Dud96] M. Dudon. *Modélisation du mouvement par Treillis Actifs et méthodes d’estimation associées : Application au codage de séquences d’images*. Thèse de doctorat de l’Université de Rennes I, 1996.
- [EY36] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1 :211–218, 1936.

- [FG87] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct point, corners and centres of circular features. In *Proceedings of the ISPRS Conference on Fast Processing of Photogrammetric Data*, pages 281–305, 1987.
- [FHK06] M. Floater, K. Hormann, and G. Kos. A general construction of barycentric coordinates over convex polygons. *Advances in Computational Mathematics*, 24(1-4) :311–321, 2006.
- [Fix78] G.J. Fix. A rational finite element basis (Eugene L. Wachspress). *Society for Industrial and Applied Mathematics Review*, 20(1) :195–196, 1978.
- [Flo97] H. Le Floch. *Échantillonnage irrégulier et interpolation : application à la représentation d’images fixes et animées*. Thèse de doctorat de l’Université de Rennes I, 1997.
- [Flo03] M. Floater. Mean value coordinates. *Computer Aided Geometric Design*, 20(1) :19–27, 2003.
- [FM91] C.-S. Fuh and P. Maragos. Affine models for image matching and motion detection. *International Conference on Acoustics, Speech, and Signal Processing*, 4 :2709–2412, 1991.
- [GB97] P.-L. George and H. Borouchaki. *Triangulation de Delaunay et maillage : application aux éléments finis*. Editions Hermes, 1997.
- [GBBS07] V. Gay-Bellile, A. Bartoli, and P. Sayd. Deformable surface augmentation in spite of self-occlusions. *IEEE International Symposium on Mixed and Augmented Reality*, 2007.
- [Geo91] P.-L. George. *Génération automatique de maillages : applications aux éléments finis*. MASSON, 1991.
- [GKS76] G.H. Golub, V. Klema, and G.W. Stewart. Rank degeneracy and least squares problems. Technical report, STAN-CS-Computer Science Department, Stanford University, 1976.
- [GL96] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press (3rd Edition), 1996.
- [GVPG03] P.F. Gabriel, J.G. Verly, J.H. Piater, and A. Genon. The state of the art in multiple object tracking under occlusion in video sequences. *Advanced Concepts for Intelligent Vision Systems*, 2003.
- [Har71] R.L. Hardy. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, 76 :1905–1915, 1971.
- [Har84] R. Haralick. Digital step edges from zero crossing of second directional derivatives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(1) :58–68, 1984.
- [HB98] G.D. Hager and P.N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 20(10) :1025–1039, 1998.
- [Hor86] B.K.P. Horn. *Robot Vision*. MIT Press, 1986.
- [HS88] C. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [HS93a] R. Haralick and L. Shapiro. *Computer and robot vision - volume I*. Addison-Wesley, 1993.

- [HS93b] R. Haralick and L. Shapiro. *Computer and robot vision - volume II*. Addison-Wesley, 1993.
- [HU98] J.B. Hiriart-Urruty. *Optimisation et analyse convexe*. Puf, 1998.
- [HW81] R. Haralick and L. Watson. A facet model for image data. *Computer Graphics and Image Processing*, 15 :113–129, 1981.
- [Jaa02] T.S. Jaakkola. Machine learning : lecture 4. Technical report, MIT Computer Science and Artificial Intelligence Laboratory, 2002.
- [JD02] F. Jurie and M. Dhome. Hyperplane approximation for template matching. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24(7), pages 996–1000, 2002.
- [JFM01] A. Jepson, D. Fleet, and T. Maraghi. Robust online appearance models for visual tracking. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages I :415–422, 2001.
- [JSR05] A. Jacquot, P. Sturm, and O. Ruch. Adaptive tracking of non-rigid objects based on color histograms and automatic parameter selection. In *Proceedings of the IEEE Workshop on Motion and Video Computing*, pages 103–109. IEEE Computer Society, 2005.
- [KH03] T. Kaneko and O. Hori. Feature selection for reliable tracking using template matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 796–802, 2003.
- [KR82] L. Kitchen and A. Rosenfeld. Gray-level corner detection. *IEEE in Pattern Recognition letters*, 1 :95–102, 1982.
- [LB00] V. Lepetit and M.-O. Berger. A semi-automatic method for resolving occlusion in augmented reality. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 225–230, 2000.
- [Len04] J. Lenoir. *Modèle déformable 1D pour la simulation physique temps-réel*. Thèse de doctorat de l'Université des Sciences et Technologies de Lille, 2004.
- [LF95] H. Li and R. Forchheimer. A transformed block-based motion compensation technique. *IEEE Transactions on Communications*, 43 :1673–1676, 1995.
- [LG01] T. Lindeberg and J. Gårding. Shape-adapted smoothing in estimation of 3-D shape cues from affine distortions of local 2-D brightness structure. Technical report, KTH Kungliga Tekniska högskolan - Royal Institute of Technology, 2001.
- [Lin98] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2) :79–116, 1998.
- [LK81] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conferences on Artificial Intelligence*, pages 674–679, 1981.
- [Low04] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60 :91–110, 2004.
- [Lue69] D. Luenberger. *Optimization by vector space methods*. John Wiley and Sons, 1969.
- [Mar00] G. Marquant. *Représentation par maillage adaptatif déformable pour la manipulation et la communication d'objets vidéo*. Thèse de doctorat de l'Université de Rennes I, 2000.

- [MB04] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2) :135–164, November 2004.
- [Mey02] M. Meyer. Generalized barycentric coordinates on irregular polygons. *Journal of Graphics Tools*, 7 :13–22, 2002.
- [MGH80] J.J. Moré, B.S. Garbow, and K.E. Hillstom. User guide for MINPACK-1. Technical Report ALN-80-74, Argonne national laboratory, 1980.
- [Mor77] H. Moravec. Techniques towards automatic visual obstacle avoidance. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, page 584, 1977.
- [MPL00] G. Marquant, S. Pateux, and C. Labit. Mesh-based scalable video coding with rate-distortion optimization. In *Visual Communications and Image Processing*, pages 967–976, 2000.
- [MS04] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 20(1) :63–86, 2004.
- [MTS<sup>+</sup>05] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schafalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1-2) :43–72, 2005.
- [NCH93] J. Nieweglowski, T. G. Campbell, and P. Haavisto. A novel video coding scheme based on temporal prediction using digital image warping. *IEEE Transactions on Consumer Electronics*, 39(3) :141–150, 1993.
- [NH94] Y. Nakaya and H. Harashima. Motion compensation based on spatial transformations. *IEEE Transactions on Circuits Systems for Video Technology*, 4(3) :339–356, 366–367, June 1994.
- [NL93] H. Nicolas and C. Labit. Motion and illumination variation estimation using a hierarchy of models : Application to image sequence coding. Technical Report RR-2043, INRIA Rennes, 1993.
- [NSB07] N. Noury, F. Sur, and M.-O. Berger. Fundamental matrix estimation without prior match. In *Proceedings of International Conference on Image Processing*, pages 513–516, 2007.
- [NW06] J. Nocedal and J. Wright. *Numerical Optimization*. Springer, second edition, 2006.
- [OK90] M. Okutomi and T. Kanade. A locally adaptive window for signal matching. In *International Conference on Computer Vision*, pages 190–199, 1990.
- [PCM04] P. Parisot, V. Charvillat, and G. Morin. Compensation de mouvement par maillage : apprentissage et maintien de la connectivité. In *Proceedings of COMPRESSION et REPRÉSENTATION des SIGNAUX AUDIOVISUELS*, pages 161–164. France-Telecom et Enic, 2004.
- [PCM05a] P. Parisot, V. Charvillat, and G. Morin. Non-rigid tracking with 2-D meshes. In *Advanced Concepts for Intelligent Vision Systems*, pages 579–586. Springer Verlag LNCS, 2005.
- [PCM05b] P. Parisot, V. Charvillat, and G. Morin. Suivi par triangulation d’objets déformables : méthode par apprentissage. In *ORASIS*, pages 1–9, <http://www.lasmea.univ-bpclermont.fr>, 2005. LASMEA - Université de Clermont-Ferrand.

- [PCT07] P. Parisot, V. Charvillat, and B. Thiesse. Plus de qualité pour des sous-ensembles de points d'intérêt, au coeur d'algorithmes de suivi fondés sur un modèle d'apparence. In *ORASIS*, page (support électronique). LSIIT - Université de Strasbourg, 2007.
- [PMC04] P. Parisot, G. Morin, and V. Charvillat. Transformations locales d'une triangulations 2-D pour la compensation de mouvement par maillages. In *Groupe de Travail en Modélisation Géométrique*, pages 113–122. LIRIS, 2004.
- [PTC07] P. Parisot, B. Thiesse, and V. Charvillat. Selection of reliable feature subsets for appearance-based tracking. In *International Conference on Signal-Image Technology and Internet Based Systems*, pages 891–898. IEEE Computer Society, 2007.
- [RD99a] P.J. Rousseeuw and K. Van Driessen. Computing LTS regression for large data sets. Technical report, Institute of Mathematical Statistics Bulletin, 1999.
- [RD99b] P.J. Rousseeuw and K. Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41, 1999.
- [RDO98a] E. Ramis, C. Deschamps, and J. Odoux. *Cours de mathématiques, 3. topologie et éléments d'analyse*. DUNOD, 1998.
- [RDO98b] E. Ramis, C. Deschamps, and J. Odoux. *Cours de mathématiques, 5. application de l'analyse à la géométrie*. DUNOD, 1998.
- [RDO01a] E. Ramis, C. Deschamps, and J. Odoux. *Cours de mathématiques, 1. algèbre*. DUNOD, seconde edition, 2001.
- [RDO01b] E. Ramis, C. Deschamps, and J. Odoux. *Cours de mathématiques, 2. algèbre et applications à la géométrie*. DUNOD, seconde edition, 2001.
- [RH95] S. J. Reeves and R. Hezar. Selection of observations in magnetic resonance spectroscopic imaging. In *Proceedings of International Conference on Image Processing*, pages 641–644, 1995.
- [RH97] P.J. Rousseeuw and M. Hubert. Recent developments in PROGRESS. In *L1-Statistical Procedures and Related Topics*, 1997.
- [RK76] A. Rosenfeld and A.C. Kak. *Digital Picture Processing*. Academic Press, 1976.
- [RL87] P.J. Rousseeuw and A.M. Leroy. *Robust regression and outlier detection*. Wiley Series in Probability and Mathematical Statistics, 1987.
- [Rob03] P. Robert. *Le Petit Robert*. Edition Dictionnaires Le Robert, 2003.
- [SB95] M. Smith and J.M. Brady. SUSAN - a new approach to low level image processing. Technical Report TR95SMS1c, Chertsey, Surrey, UK, 1995.
- [SB97] M. Smith and J.M. Brady. SUSAN - a new approach to low level image processing. *International Journal of Computer Vision*, 23(1) :45–78, 1997.
- [She] J.R. Shewchuk. Triangle : A two dimensional quality mesh generator and Delaunay triangulator. version 1.6. <http://www-2.cs.cmu.edu/~quake/triangle.html>.
- [SMB00] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2) :151–172, 2000.
- [SS97] H. Shum and R. Szeliski. Panoramic image mosaics. Technical Report MSR-TR-97-23, Microsoft Research, 1997.

- [ST94] J. Shi and C. Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
- [TAZ05] S. Tabbone, L. Alonso, and D. Ziou. Behavior of the laplacian of gaussian extrema. *Journal of Mathematical Imaging and Vision*, 23(1), 2005.
- [TH98] M. Trajkovic and M. Hedley. Fast corner detection. *Image and Vision Computing*, 16 :75–87, 1998.
- [TK91] C. Tomasi and T. Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, 1991.
- [Tok98] C. Toklu. *Object-based Digital Video Processing using 2-D Meshes*. PhD Thesis University of Rochester, USA, 1998.
- [Tri04] B. Triggs. Detecting keypoints with stable position, orientation and scale under illumination changes. In *European Conference on Computer Vision*, pages 100–113, 2004.
- [Urb03] M. Urban. Harris interest operator. *lecture note* : [http://cmp.felk.cvut.cz/cmp/courses/dzo/resources/lecture\\_harris\\_urban.pdf](http://cmp.felk.cvut.cz/cmp/courses/dzo/resources/lecture_harris_urban.pdf), 2003.
- [WB95] H. Wang and M. Brady. Real-time corner detection algorithm for motion estimation. *Image and Vision Computing*, 13(9) :695–703, 1995.
- [WD04] A. Warusfel and C. Deschamps. *Mathématiques tout-en-un 2e année MP - Cours et exercices corrigés*. DUNOD, 2004.
- [WL94] Y. Wang and O. Lee. Active mesh : A feature seeking and tracking image sequence representation scheme. *IEEE Transactions on Image Processing*, 3(5) :610–624, 1994.
- [ZdH02] Z. Zivkovic and F. Van der Heijden. Better features to track by estimating the tracking convergence region. In *International Conference on Pattern Recognition*, pages 635–638, 2002.
- [ZdH04] Z. Zivkovic and F. Van der Heijden. Improving the selection of feature points for tracking. *Pattern Analysis and Applications*, 7(2) :144–150, 2004.
- [ZWT99] Z. Zheng, H. Wang, and E. Teoh. Analysis of gray level corner detection. *IEEE in Pattern Recognition letters*, 20 :149–162, 1999.